CGATATCATT

#### 1/82

#### PROTEIN GAGCTGAACG GCCACTATGC GAAAAAGGCA AAAGCCACTT AATCTGTTTT ACAGCGTTCT CTTAATACTA GTACAAACCC ACAATAAAAT ATGACAAACA TTAAAAAATA TCATCTTTCA CATCTTTCAT ACATGCCCTG AACTAACCTT AGGAGAAAAT ATGCTTTGGT WEIGHT AAACGCCTGA ACTTAGCGTT TCTATTCCAC ACACGGCACA TTCATCTTTC AATGAAGAGG CCATTCCACA TCTTTCATCT TTCATCTTTC ATCTTTCATC TTTCATCTTT TGCAAATATT CTTTCATCTT MOLECULAR TGCTCGCATG AAAGTGCGTC GAGGGCCAAG ATAAAGTAAT TTAATTGTTC TATATCGTCT CAAATTCAGC AGGTGTAACA TGGATGTAGT GAATTGGCAC GGGGTTGTGA GCAGTCTATA ATGGTATAAT TCATCTTTCA TCTTTCATCT HIGH OF. TACTATCTTT TTACAAGGAA FIG. 1A. DNA SEQUENCE GGAAGGGAGG CACCTTTTTT GCCATATAAA (HMM1 Н GCGAAAAACC TCCGCTATGT ACAATTACAA STATAAATCC CTTTCATCTT ATGAACCGAG AACGCAAATG ATGAACAAGC TGCTGTGTCT AGCAAGCGGC 401 451 501 551 57 151 201 251 301 351 101

AGTTTTACA TGTTACATCT AACCAAATCT TAATAAAACC ATTATCCGCA ACAGTGTTGA GAAATGGTGC TATTCAACCG CGACCAAAAT AATTTAACAT AACTCCGCCG AGAAAACAAC AATTGGAAAC 701 601 651

AAGTAGATGG

## FIG. 1B

GTTAATTGAC	GCGATATTGC	ATTGTGTGGG	CGGACGCGCT	AAGAAAAAGG	501
GTATCAGGCA	AACCATCAAT	AAAAAGGCTC	ACCTCTTTAG	AGCAAAGAAA	451
GCATTCAATŤ	GGTAAAAAGG	GCGCGGCGAA	GCGGTGACGA	ACTTACCTTG	.401
AGGGGGAGAA	CAGGTAAAGA	ATCGACCTTT	AGGTGCAGTT	CATTAAAAAC	1351
GATAAAGTCA	GATTACAGGC	GCAAGCTGAT	GCTAAAGGCG	AAATCAGCAA	1301
TTTCCGCTCA	GGCGGTGTAA	AGCGGAAATT	AAGAGGGTGA	CTTTCCGCCA	1251
CAATATTGTT	ATAAAAGCGG	GTAAGCAAAG	TGCTGATTCT	GTAAACTTTC	1201
CGAAACCAAG	TGCCACTATT	ATGTCCGTGC	GGTAACATTA	TGCCAAAGGC	1151
GCGATATTT	GTCAATCTGG	AAATGAAGCG	CCGCGCCTGA	TACAGCATTG	1101
AACCATTACT	TAATAAACCC	ATCAGCGATA	AAAAATCACC	TCGCAGGGCA	1051
ATTTCTTTAC	TGGTGGCAGC	TTAGCGTAAA	GAGGGTGTGA	AGTGAAAAAC	1001
TTGGTGGCAA	GTAAATCTTA	AGACGGCAGT	CTGTCGGTAA	GGTTTAATTA	951
TGTGAATCAC	TCGCTGAAAT	GATAAAGCGC	GCAAACCAAA	TCACCTTCGA	901
GCGCGTAATT		TTTCTAACGA	ACGCTAGACA	TACGGCTTCT	851
CTAATGGCTT		TAAAGACGCA	. TCACAATAGG	CCAAATGGTA	801
TTTAATCAAC	GACAAGTCTT	GATTCTAACG	, AGGGATTTTA	CCCAATTAAA	751

# FIG.1C.

ACGCTCAAGG TAGTGGTGAT ATCGCTAAAA CCGGTGGTTT GCAATTGTTG TAATGCAGAA CGGGATCCGG ACATTAACAA TAACATCACT CCAATGGCAG GTCGGAGCGG TGGCGGCGTT GAGATTAACA GTGCAAACTT AACAATTTAC ATCTCACTCG GGGCGCAAGG AAAGGAAGCA TCAAAAAGGT GACTGCAATT TTTGAAGGGA ACCTAAAAAT ATTTAACCTC CAAAGACAAT ATGTATCTAT AACGAAACAA AGAAAAGACA TTCAGAAGAC GATGAATACA GTACCTTTGT CAATAGCTCC ATTAATTTAT CGCCTTTGAG CCTCAGGCAA TCTAAACGGC ACTGGCAGCG CACAAATAAA CAAAGGACGC ACTTACTGGA CAATGGTTTT GACCCGGATA ATTTATTCAT CTAAAAAAG CACCGGTGAT GATACCAGAG TCATAAAAAT ATTACAGCTA AACAAGATAT GGGACTATTA AATACGCTAT GTGAACATCT TCGGGGCATG GTGGTTGTTA GAATAGTGCC AGCACCCCAA GCAGCAATAC ACACAACTCT TGAGAGTATA GCATCTATGT TGGAGTGAGG GGGTTGATGT TACAGGTCAA ATAATGTCTC AGAACCAATA TTCAGGGAAA ATGATAAATT GGCAATATTA TGTGGAGACG ACAGCAGGAC ACGCCAAAGA GCTAATCAAC CTTAACTCTT ACGATATTAC TCAGGCGGCT TAACATAAAC ACCAAGTCAT TTTAGATTTA CACCACTAAA CTTTAAATAT GAAAGTGGAT 1601 51 1651 1701 1751 1851 1801 1901 1951 2001 2051 2101 2151 2201 2301 2251 S

## FIG. 1D.

2351	CTTAAATGTT	TCCGAGAGTG	GCGAGTTTAA	CCTCACTATT	GACTCCAGAG	
2401	GAAGCGATAG	TGCAGGCACA	CTTACCCAGC	CTTATAATTT		
2451	TCATTCAACA	. AAGACACTAC	CTTTAATGTT	GAACGAAATG		
2501	CTTTGACATC	AAGGCACCAA		ТААСТАТТСТ	AGTTTC A ATT	
2551	ACGCATCATT	TAATGGAAAC	•	CGGGAGGGGG		
2601	TTCACACTTC	TCGCCTCATC		CAAACCCCCG		
2651	AAATTCTAAA	TACTTTAATG		GTCAAGTTTA		
2701	CTTCAGGCTC	AACAAAAACT		TAGAGAGAGA	4/2 « ВВВО « «БББ	
2751	AATGCCACCG	GAGGCAACAT	AACACTTTTG	CAAGTTGAAG		
2801	AATGATTGGT	AAAGGCATTG	TAGCCAAAAA	ТАGCСААДАД ДАДСАТААСС		
2851	GTAACATCAC	CTTTGGCTCC				
2901				INTORCHORANI	CGAAGGCAA'I'	
	GITACIAICA	ATAACAACGC	TAACGTCACT	CTTATCGGTT	CGGATTTTGA	
2951	CAACCATCAA	CAACCATCAA AAACCTTTAA	CTATTAAAAA	AGATGTCATC	ATTAATAGCG	
3001	GCAACCTTAC	CGCTGGAGGC	AATATTGTCA	ATATAGCCGG	AAATCTTACC	
051	GTTGAAAGTA	ACGCTAATTT	CAAAGCTATC	ACAAATTTCA		
101	AGGCGGCTTG	TTTGACAACA			ATTGCCAAAG	
151	GAGGGCTCG	CTTTAAAGAC ATTGATAATT			AAGCATCACC	
		_				

## FIG. 1E.

		1	<b>.</b>					5/	82								
		ATAGCGGCA ATATAACCAA	ACTGAAATGC	GATTTCTTCT	GTGTTGATGG	ACCATTAAAA	TTTCAATAAA	GTAACACCAA	CAGGTTAAAG	CAGCAAAGTG	ACAATAATGC	AATATTACTT	TACCACTAAÀ	TAACCGCTCA	TCTGTAACAC	GGGCAACACC	AGGCTCTAC
			AGGTAGTGAT	GTAATCTCAC		TGCCAATCTA	ATATTTCAGG	TTAACTATTG	AACCTTTAAC CAGGTTAAAG	TGACACTACA	GATAGCAGTG					GCAATATTTC G	ACCACTTTGG CAGGCTCTAC
	GCTCCACTTA CCGCACTAGE				ACAGATAACA		TGGTPGTC1AA			GOICACAAGG			T'C'IGCGACAA (	AACCACTGGT AACGTGGAGA		CIIGCIGI'AA (	CGGTGCATTA A
										GTAGTAATAA		AGTENCYAN AAAATGTAAC					
•	ACCAACTCCA	TAAAAACGGT	AAATTGGCGG	GACAAAATCA	GGAGAATTCC	CCAAAGAATT	GCAGAGATTA	TAGTGCTGAT	ATTCAAAAAT	GAAACATCCG	CGGCTTAACT	CTCACAAAGC					
	3201	3251	3301	3351	3401	3451	3501	3551	3601	3651	3701	3751	3801	3851	3901	951	

#### TG.1F

TACACAAAT	TTACAG'I'CGA	AATACAA		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	 
	( ) ( ) ( ) ( ) ( )	بر 3 د د 3 د د 3		い中心中ないの中中の	801
GGAGTAAGTG ·	AGCTAAACTT	GAGAAGCGTT	GATGAAGAAA	AGATTTATCT	751
AGAAGGTAAA	CGCATCCTTG	TGAAGCGAAA	ATGAAGTAAT	GCAAGCGTAG	701
ACCGGGTATA	AATACATTCA	ATTGATGTGA	AGGCGTTAAA	TACTGTTAAA	.651
ATAAACACCG	AAAAAACGGT	ATATCATTTC	AATGGATTAA	AATCACAATA	601
CTGGGGATTT	GTGAACATCA	CTCAAGCAGA	TCGCGACAAC	GGCAGCGTAA	1551
AAATGGCTCC	CAACCAACGC	GTGGTAAATG	TAACCACACA	CAGCATTGGG	1501
CTAAATGGCG	AGACGCTGAG	TTAACGCAAA	ACCTTGGTTA	AACCAGCGGT	1451
ACATTAATGC	AAGGGTTCAA	AACTACCGTG	CAGGCACTTT	CTAAATACTA	1401
CAATGTGACA	TTAATGCCGC	GCAGGAAGTA	TGGTAGCGTT	CAGCTCAGGA	1351
GTAAATCTTT	CAAGGGTCAG	TTACTTCAGC	AGTTCACACA	TACCGAAGCT	4301
GCAAATTAAC	ACATCATCGG	AACCTTAACT	AAGGAGCTGC	AATGCGACAG	4251
CGCAGAAATT	TTGGGAATGG	GATTTAACAG	AAACGCTGGC	ATGTTACGGC	4201
AATACGGTAA	GATTTCCGGT	TTGGTGGTAC	ACAGGTACAA	AACAAGTGCA	4151
AGGCTAACGT	ACAACAGGCG	AATTAAAGCA	CCAATTCAAA	ACCACTCAAT	4101
CGAAAGTTTA	TTAAAGCAAC	ACAGTAGAGG	TTCTGGTGGC	GCGGTACGAT	4051
GGCGATATCG	AAGTCAATCA	TAACCACTTC	ACCGAGAGTG	AATTAAAGGA	4001

#### 7/82

### FIG. 1G

4851	GAATTTGCAA	CCAGACCATT	GAATTTGCAA CCAGACCATT AAGTCGAATA GTGATTTCTG AAGGCAGGGC	GTGATTTCTG	AAGGCAGGGC
4901	GTGTTTCTCA	AACAGTGATG	GTGTTTCTCA AACAGTGATG GCGCGACGGT GTGCGTTAAT ATCGCTGATA	GTGCGTTAAT	ATCGCTGATA
4951	ACGGGCGGTA	GCGGTCAGTA	GGTA GCGGTCAGTA ATTGACAAGG TAGATTTCAT CCTGCAATGA	TAGATTTCAT	A D T A A D D D D
5001	AGTCATTTTA	TTTTCGTATT	AGTCATTTTA TTTTCGTATT ATTTACTGTG TGGGTTAAAG TTCACTACTACTACTACTACTACTACTACTACTACTACTA	TGGGTTA A A C	
5051	GCTTTACCCA	TCTTGTAAAA	CCCA TCTTGTAAAA AATTACGAG AATACAATAAAAAAAAAA		
5101	ACAGGTTATATT ATTATO			441440ain	AGIALLITA

#### 8/82

#### WEIGHT MOLECULAR HIGH O FJ SEQUENCE FIG. 2A. AMINO ACID PROTEIN

KVRHLALKPL DSNGQVFLIN ISDIINPTIT DFDNVSINAE GTITSGNOKG IIRNSVDAII IDLSGKEGGE IVWGDIALID LKKGTFVNIT DKALAEIVNH VSKDKSGNIV DTRGANLTIY VNISMVLPKN LTQPYNLNGI ISVSGGGSVD EKGSEKPARM FEGTLNISGK DSRGSDSAGT ARNFTFEQTK ISLLAGOKIT DKVTLKTGAV EINNDITTGD KGSNQVITGQ RNQGKLSADS VSGKEKGGRA TLTNTTLESI ATMOVDGNKT NQISQLKGIL AIVDAKEWLL SLNYASFNGN ELARGCDHST NSAVFNRVTS TLDISNENIK EGVISVNGGS LQGMDVVHGT GNINVRAATI AKGGKLMITG TSLEKGSTIN STPKRNKEKT WSEGRSGGGV ITAKQDIAFE RTNKYAITNK KAPIGINKYS SGHDLFIKDN SESGEFNLTI KRLNALVAVS SIPQSVLASG EMVQFLQENN IINTNGFTAS VNLIGGKVKN VNLGDIFAKG GGVISAQNQQ TGSGLOFTTK GKNGIQLAKK IAKTGGFVET INLSNGSLTL ISLGAQGNIN TYWNLTSLNV DEYTGSGNSA ERNARVNFDI MNKIYRLKFS GLITVGKDGS FRFNNVSLNG NWKQFNIDQN PNGITIGKDA YSIAAPENEA TYLGGDERGE GNINAQGSGD SAMLLSLGVT TAGRSNTSED LSAKEGEAEI ANORIYVNSS SGGWVDVHKN ESGYDKFKGR SFNKDTTFNV 51 101 151 201 301 251 351 401 451 501 551 601 701 651

#### FIG. 2B

Gr ひ1 氏 大 し し 「 」	RKAVTEIEGN	NIVNIAGNI,T	IDNSKNLSIT	OKEGNLTISS	QDLNISGFNK		SATSGEITTK 8	LAVSNISGNT	TVEVKATESL	DLTVGNGAEI	AGSINAANVT	VVNATNANGS	IDVKYIQPGI	NOTITVDTON	
				_											
RFKTSGSTKT	FEGGNITFGS	INSGNLTAGG	IAKGGARFKD	TEMQIGGDVS	TIKTKELKLT	QVKDSKISAD	NITSHKAVSI	SVTLTATEGA	GDIGGTISGG	NTVNVTANAG	VNLSAQDGSV	LNGAALGNHT	INTVLLKGVK	GVSAVRFIEP	IADNGR
YFNVSTGSSL	KGIVAKKNIT	KPLTIKKDVI	FDNKGNSNIS	DLNITNEGSD	DSDATNNANL	GTNAKKVTFN	IDAKNVTVNN	ILGGIESSSG	TESVTTSSQS	TGTIGGTISG	SSHITSAKGQ	TLVINAKDAE	NGLNIISKNG	DEEREALAKL	NSDGATVCVN
Q'1'FGVVINSK	QVEGTDGMIG	LIGSDFDNHQ	TNFTFNVGGL	ISGNITNKNG	IKAGVDGENS	LTIGNTNSAD	DSSDNNAGLT	NVEITAQTGS	TTLAGSTIKG	TTGEANVTSA	TSSGKLTTEA	KGSNINATSG	VNITGDLITI	RILEKVKDLS	VISEGRACFS
r i llassassiv	NATGGNITLL	VTINNNANVT	VESNANFKAI	TNSSSTYRTI	DKINITKQIT	AEITAKDGSD	ETSGSNNNTE	TGTTINATTG	VTVTANSGAL	TTQSNSKIKA	NATEGAATLT	LNTTGTLTTV	GSVIATTSSR	ASVDEVIEAK	EFATRPLSRI
T C /	801	851	901	951	001	051	101	151	201.	251	301	351	101	151	01

### FIG.3A.

DNA SEQUENCE OF HIGH MOLECULAR WEIGHT

PROTEIN II (HMW2)

10/82 AATAAATCAA GATTTTTGTG ATGACAAACA TTAAAAAAT GGAGCTGAAC CACATGAAAT TAGGAGAAAA TTCATCTTTA TCATCTTTCA AATGCTTTGG ACCATTCCAC AGAAAAGGC TAAAGCCACT CAATCTGTTT AGCCACTATG ACGCTATCAT TAACCAAATC CAGTTTTTAC CTTTCATCTT GAATGAAGAG CAAACGCCTG TGCAAATATT TCTTTCATCT TTTCATCTTT CAACTAACCT CACTTAGCGT ATCTATTCCA TACACGGCAC AACAGTGTTG TGAAATGGTG CAACTCCGCC GTATTCAACC GTGTTACATC GCAGTCTATA AATGGTATAA CATCTTTCAT GATGAACCGA GGGAAGGGAG GGAGGGGCAA TIGCIGIGIC IGAATIGGCA CGGGGTIGIG TTACTATCTT TAGGTGTAAC TAGGTGTAAC ATCTTTCATC TTTAATTGTT TCAAATTCAG CTTACAAGGA ATGGATGTAG CATTATCCGC TCGACCAAAA AGATAATAAA CGCCATATAA TTTCATCTTT CACCTTTTTT TTCATCTTTC GATAAAGTAA ATATATCGTC GTAÄTAAAAC TTACTATCTT CAATTTAACA TAAATATACA ACAATTACAA AGTATAAATC TCTTTCATCT TAGCAAGCGG ATCTTTCATC SAACGCAAAT TATGAACAAG TTCCGCTATG AAGAAAACAA TTCCGCTATG CAAGTAGATG TAATTGGAAA  $\leftarrow$ 51 101 151 201 251 301 51 401 451 501 551 651 601 701

#### FIG.3B

						1	1/82								
TTTTAATCAA	ACTAATGGCT	GGCGCGTAAT	TTGTGAATCA	ATTGGTGGCA	CATTTCTTTA	CAACCATTAC	GGCGATATTT	TCGAAACCAA	GCAATATTGT	ATTTCCGCTC	CGATAAAGTC	AAGGGGGAGA	GGCATTCAAT	TGTATCAGGC	CGTTAATTGA
GGACAAGTCT	AATTATTAAC	AAAACATCAA	CTCGCTGAAA	TGTAAATCTT	ATGGTGGCAG	ATAATAAACC	GGTCAATCTG	CTGCCACTAT	GATAAAAGCG	TGGCGGTGTA	TGATTACAGG	TCAGGTAAAG	AGGTAAAAAC	CAACCATCAA	GGCGATATTG
AGATTCTAAC	GTAAAGACGC	ATTTCTAACG	AGATAAAGCG	AAGACGGCAG	ATTAGCGTAA	CATCAGCGAT	AAAATGAAGC	AATGTCCGTG	TGTAAGCAAA	AAGCGGAAAT	GGCAAGCTGA	TATCGACCTT	AGCGCGGCGA	GAAAAAGGCT	TATTGTGTGG
AAGGGATTTT	ATCACAATAG	TACGCTAGAC	AGCAAACCAA	ACTGTCGGTA	CGAGGGTGTG	AAAAAATCAC	GCCGCGCCTG	CGGTAACATT	CTGCTGATTC	AAAGAGGGTG	AGCTAAAGGC	CAGGTGCAGT	GGCGGTGACG	AACCTCTTTA	GCGGACGCGC
TCCCAATTAA	CCCAAATGGT	TTACGGCTTC	TTCACCTTCG	CGGTTTAATT	AAGTGAAAAA	CTCGCAGGGC	TTACAGCATT	TTGCCAAAGG	GGTAAACTTT	TCTTTCCGCC	AAAATCAGCA	ACATTAAAAA	AACTTACCTT	TAGCAAAGAA	AAAGAAAAG
751	801	851	901	951	1001	1051	1101	1151	1201	1251	1301	1351	1401	1451	1501

## FIG.3C.

1551	CGGCAATATT		AACGCTCAAG GTAGTGGTGA	TATCGCTAAA	TATCGCTAAA ACCGGTGGTT	
1601	TTGTGGAGAC	ATCGGGGCAT	TATTTATCCA	TTGACAGCAA	TGCAATTGTT	
1651	AAAACAAAAG	AGTGGTTGCT	AGACCCTGAT	GATGTAACAA	TTGAAGCCGA	
1701	AGACCCCCTT	CGCAATAATA	CCGGTATAAA	TGATGAATTC	CCAACAGGCA	
1751	CCGGTGAAGC		AAGCGACCCT AAAAAAATA GCGAACTCAA AACAACGCTA	GCGAACTCAA	AACAACGCTA	
1801	ACCAATACAA		CTATTTCAAATTATCTGAAA AACGCCTGGA CAATGAATAT	AACGCCTGGA	CAATGAATAT	
1851	AACGGCATCA	AGAAAACTTA	CCGTTAATAG	CTCAATCAAC	ATCGGAAGCA	1
1901	ACTCCCACTT	AATTCTCCAT	AGTAAAGGTC	AGCGTGGCGG	AGGCGTTCAG N	2/82
1951	ATTGATGGAG	ATATTACTTC	TAAAGGCGGA	AATTTAACCA	TTTATTCTGG	
2001	CGGATGGGTT	GATGTTCATA	AAAATATTAC	GCTTGATCAG	GGTTTTTAA	
2051	ATATTACCGC	CGCTTCCGTA	GCTTTTGAAG	GTGGAAATAA	CAAAGCACGC	
2101	GACGCGGCAA	ATGCTAAAAT	TGTCGCCCAG	GGCACTGTAA	CCATTACAGG	
2151	AGAGGGAAAA	GATTTCAGGG	CTAACAACGT	ATCTTTAAAC GGAACGGGTA	GGAACGGGTA	
2201	AAGGTCTGAA	TATCATTTCA	TCAGTGAATA	ATTTAACCCA CAATCTTAGT	CAATCTTAGT	
2251	GGCACAATTA	ACATATCTGG	GAATATAACA	ATTAACCAAA	CTACGAGAAA	
2301	GAACACCTCG	TATTGGCAAA	CCAGCCATGA	TTCGCACTGG	AACGTCAGTG	
2351	CTCTTAATCT		AGAGACAGGC GCAAATTTTA CCTTTATTAA ATACATTTCA	CCTTTATTAA	ATACATTTCA	

#### FIG.3D

						1	3/82								
CAGGGGTGAA	GAAGGAGCGA	AAGCAAACCT	GGGCTCTGT	GAGTTAAAAA	AAATTCCCAT	CCATAAATGC	TTTTATGACG	CATTCTGGGC	TTACGGGGAA	AATAACGCCC	CAGCTTGCTC	TTAAAGGCAA	AGAGATACCC	AATTAATATA	GTGATTTAAA
AGAAGCTCTG	CAATCTCAAA	ACATGAACAC	GCCACTGGTG	CAGAGGGGCT	ATTTTACCTT	AAAGACTTAA	GAAAGATGAT	ACAACATATC	AGCAGCAGCA	GCTAGAAGCC	TAAAACTTGG	AATGCAGATA	AGGAAAGACT	GCACTGCCGA	ACCAATGATG
AACACAGTAT	ACATGTCATT	CCAAACGAGA	CAATATCACA	ACCATTCTGG	AACGGCGCTA	TAAAATCAAC	TCAGACAGAC	AATTCAACCT	ACAAAACTCA	CAAATGTTAC	GATAGAGTTA	AACTGGCGAA	CCACTTTTAA	ACCAATAA†G	TGGCAATGTT
AAGGCTTAAC	GTAAATGGCA	CAAATTAAAA	GGTTTTTAGC	ATATATGCCA	TAATATCTCT	ATGACGCTTT	AATTTCAGCC	CAATGCCATC	CCCTTGGTGG	GAGAAAGCAG	AAACATAAGG	GTTTAAGTTT	TCAGAAAGCG	CGGCAATTTT	TGGTAAAACT
AGCAATAGCA	TTTTAACGGC	AAGTTAATTT	TTACCAATTC	TTTTTTGAT	TGAGTGAAAT	GTTCGCGGCG	AACCAATTCA	GGTACGCACG	GGTAATGTCA	TATTACTATC	CTAATCAGCA	GTTAATGGGA	TCTCACTATT	TAAATATCAC	ACACAAGGAG
2401	2451	2501	2551	2601	2651	2701	2751	2801	2851	2901	2951	3001	3051	3101	3151

### FIG.3E.

							14/8	2							
GGCGGAGATA	TAATGATGCT	ACCTCACGAT	AAAAAGGGTA	CAACCTAACT	TTTCAGGTTT	ACTATTGGCA	AACTTTTAAC	TGACACTAAA	AGCAATAGCG	AGTAAACAAA	AAAAGGTTAC	GCAAGTATTA	CACGGTAAGT	AAATTGAAGC	ATTGGCGGTA
AAGCATCATC	CAGACAGTAA	AAAGAAGGCA	GATAACAATC	CAAGTAATGC	GACCTAAGTA	TAGAGATTTA	CCAAAACAGT	GGTCACAATG	CGGACGTGAA	AAAATGTAGA	ACCGCGTCGG	AAATGGCAAA	TTTCCGGTAA	TCCGGCTCAA AAATTGAAGC	AACAGGTACA
GCAACCAAAG	TTAAATATTA	TATCTCGCAA	TCACCAAACA	TCAGATGCGA	ATTGACAGAA	CCAAAGAŢGG	GGTGCCGAAG	CTCTGCTGAC	GCAGCAATGG	ATTACTGCAA	AGTAAATATC	TTAACGCAAC	AGCGGTACGA	AACCACTAAA	TAACAAGTGC
CACGCTAAAC	AAAAGGAAGC	TTGGCGGCAA	AAAATTAATA	GGACTCTAGT	AAGAATTGAA	CAATAAAGCA GAGATTACAG	CGGTAACAGC	ATTCAAAAAT	AAAACATCTA	CGGCTTAACT	CTCTCAAAAC	GGCTCGACCA	AGGTGATATC	CTGGTGATTT	GAAATCGGGT GAGGCTAATG
CATTACCACT	TAATCAACAA	GAAATCCAAA	TTCTTCCGAT	TTGATGGAGA	ATTAAAACCA	CAATAAAGCA	ACAGTAATGA	AATGTTAAAG	TAGCAAAGTG	ACAACGATAC	GATATTACTT	CACCACAGCA	CAACCAAAAC	GTTAGCGCGA	GAAATCGGGT
3201	3251	3301	3351	3401	3451	3501	3551	3601	3651	3701	3751	3801	3851	3901	3951

## FIG.3F.

							15/	82								
CGATTTAACA					ATTAACGCAA	AGAAGTGAAT	CCTCAAGCAG	AATATCATTT	AATTGAGGTG	TTGAAGCGAA	AGAGAAACAT	AAATAATACA	CAAGTCAAGT	GGCGCACGAG	AATTGACAAG.	TATTTACTGT
CAAACGCTGG	GAAGGAGCTG	CGGTTCTAGC	ATGGTAGCAT	ACAGGCACCT	CACCTTGGTT	GTGATAGTAC	ACTGCGGCAA	AAATGGGTTA	GAGGCAAGGA	GAAGAAGTAA	TGATGAAGAA	TTGTTGAGCC	ACCAGACCGT	AAGTGGTAAT	CGTAGTCAGT	ATTTTCGTAT TATTTACTGT
AATGTTACGG	TAATGCGACA	CTACTGAAGC	TTGGCTCAGA	ATTAAATACT	CAACCAGCGG	GATGCATCAG	TGGTAGTGTG	TAAACACAGT	GTGCGCTTAA	AGCAAGTGTA	AAGATTTATC	GCTGTACGTT	TGAATTTACA	CGTGTTTCTC	GATGGACAGC	AAGTCATTTT
CAATTTCCGG TAATACGGTA	GCGCAGAAAT	AATACCTTGA	GGTAGACCTC	CTAATGTGAC	GATATTAAAG	AAGATGCTAA GCTAAATGGT	CAAGCGGCTC	ACTGGGGATT	TAGAAACACT	AGCCAGGTGT	GAAAAAGTAA	TGGTGTAAGT	ATACACAAAA	GAAGGTAAGG	TGTTGCTGAC	TCCTGCAATG
CAATTTCCGG	GTTGGGAATG	CGCAACAGGG	CTAAGGGTCA	ATTAATGCTG	GGCAGGCTCG	AAGATGCTAA	GCAGTCAACG	TGTGAATATC	CGAAAGATGG	AAATATATCC	ACGCGTCCTT	TAGCTAAACT	ATTACAGTCA	GATAATTTCT	TATGTACCAA	GTAGATTTCA
4001	4051	4101	4151	4201	4251	4301	4351	4401	4451	4501	1551	1601	1651	1701	1751	1801

#### 16/82

## FIG.3G.

GTGGGTTAAA GTTCAGTACG GGCTTTACCC ATCTTGTAAA AAATTACGGA GAATACAATA AAGTATTTTT AACAGGTTAT TATTATG 4851 4901

# MOLECULAR WEIGHT HIGH OF FIG. 4A. AMINO ACID SEQUENCE PROTEIN 2

## FIG. 4B.

			VADDGQP	SGNGARVCTN	IISEGKACFS	451
	EFTTRPSSQV	NNTITVNTON	GVSAVRFVEP	DEERETLAKL	RVLEKVKDLS	401
	ASVEEVIEAK	IEVKYIQPGV	RNTVRLRGKE	NGLNIISKDG	VNITGDLNTV	351
	GSVTAATSSS	EVNAVNASGS	LNGDASGDST	TLVINAKDAK	AGSDIKATSG	301
	LNTTGTLTTV	AGSINAANVT	VDLLAQNGSI	GSSITSTKGQ	ATGNTLTTEA	.251
	NATEGAATLT	DLTVGNGAEI	NTVNVTANAG	TGTIGGTISG	KSGEANVTSA	1201
	TTKSGSKIEA	TVSVSATVDL	GDISGTISGN	NGKASITTKT	TTAGSTINAT	1151
<i>-</i>	VNITASEKVT	VNKDITSLKT	GLTITAKNVE	GRESNSDNDT	SKVKTSSSNG	1101
. 0,	SADGHNVTLN	TFNNVKDSKI	GNSGAEAKTV	RDLTIGNSND	NKAEITAKDG	1051
	LTEDLSISGF	NLTIKTKELK	DSSSDATSNA	ITIKKGIDGE	SSDKINITKO	1001
	ISQKEGNLTI	NDAEIQIGGN	KGSLNITDSN	SIIGGDIINK	ITTHAKRNQR	951
	GNVTNDGDLN	INITQGVVKL	GNFTNNGTAE	GKTRDTLNIT	. LTISESATFK	901
	TGENADIKGN	SLLVNGSLSL	NIRDRVIKLG	LEANNAPNQQ	ITIEKAANVT	851
		ILGGNVTLGG	NAINSTYNIS	KDDFYDGYAR	TNSNFSLRQT	801
	KINKDLTINA	NSHVRGDDAF	NISNGANFTL	RGAELKMSEI	FFDIYANHSG	751

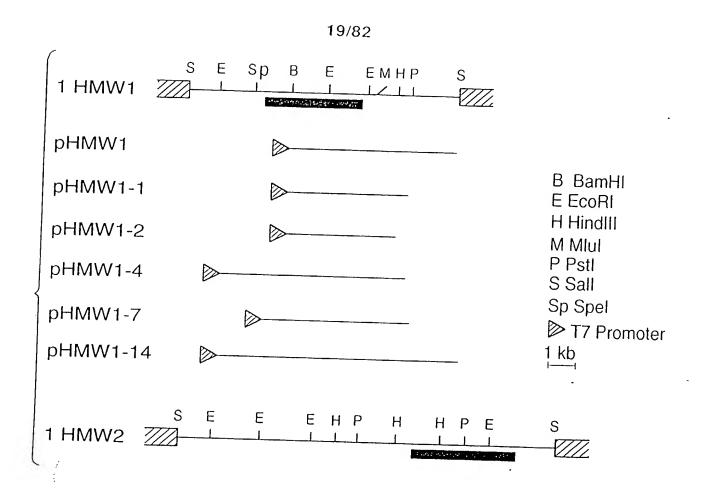
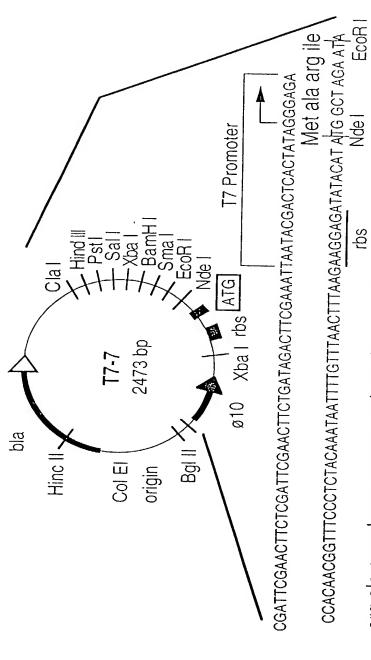


FIG.5A.





arg ala arg gly ser ser arg val asp leu gln pro lys leu ile ile asp ... cGC GCC पेंड GGA Tपेट TCT ACA GTC GAC पेंड CAG CCC AAG CTT ATC ATC ATC ATC... Smal BamHI Xbal SallPst। Hind III Clal

F16.5B.

shaded boxes indicate the locations of the structural genes. In the recombinant phage, transcription proceeds from left to right for the HMW1 gene and from right to left for the HMW2 gene. The methods used for construction of the plasmids shown are (A) Partial restriction maps of representative HMW1 and HMW2 recombinant phage and of HMW1 plasmid subclones. The described in the text. (B) Restriction map of the T7 expression vector pT7-7. This vector contains the T7 RNA polymerase promoter 女10, a ribosome - binding site (rbs), and the translational start site for the T7 gene 10 protein upstream from a

multiple cloning site (37).

### FIG.6A.

ATGACAAACA TTAAAAAATA TCATCTTTCA CATCTTTCAT ACATGAAATG GAGCTGAACG AGGAGAAAAT ATGCTTTGGT GAAAAAGGCA AAAGCCACTT GCCACTATGC CGCTATCATT AATCTGTTTT AGTTTTACA AACCAAATCT TTTAATCAAC CTTAATACTA GTACAAACCC ACAATAAAAT CTTTCATCTT GAGGGCAAG AATGAAGAGG TGCAAATATT TTTCATCTTT AAACGCCTGA TTCATCTTTC AACTAACCTT CCATTCCACA ACTTAGCGTT ACACGGCACA TCTATTCCAC ACAGTGTTGA GAAATGGTGC TGTTACATCT GACAAGTCTT ATGGTATAAT GCAGTCTATA ATCTTTCATC TCTTTCATCT TTAATTGTTC CAAATTCAGC GGGGTTGTGA AAAGTGCGTC TGGATGTAGT AGGTGTAACA ATTATCCGCA CGACCAAAAT TATTCAACCG GATTCTAACG CACCTTTTTT TCATCTTTCA GCCATATAAA TTCATCTTTC GGAAGGGAGG TATATCGTCT ATAAAGTAAT GAATTGGCAC TACTATCTTT AATTGGAAAC AATTTAACAT TGCTCGCATG TAATAAAACC TTACAAGGAA AGGGATTTTA AACTCCGCCG ACAGCGTTCT ACAATTACAA CTTTCATCTT STATAAATCC TCTTTCATCT ATGAACCGAG AACGCAAATG ATGAACAAGA TGCTGTGTCT GCGAAAAACC TCCGCTATGT AGCAAGCGGC AAGTAGATGG AGAAAACAAC CCCAATTAAA  $\leftarrow$ 51 101 201 151 251 301 351 401 451 501 551 601 651 701 751

## FIG. 6B.

GCAATTGTTG	CAAAGACAAT GCAATTGTTG	ATTTATTCAT	TCGGGGCATG	TGTGGAGACG	601
CCGGTGGTTT	ATCGCTAAAA	TAGTGGTGAT	ACGCTCAAGG	GGCAATATTA	551
GTTAATTGAC	GCGATATTGC	ATTGTGTGGG	CGGACGCGCT	AAGAAAAAGG	501
GTATCAGGCA	AACCATCAAT	AAAAAGGCTC	ACCTCTTTAG	AGCAAAGAAA	451
GCATTCAATT	GGTAAAAACG	GCGCGGCGAA	GCGGTGACGA	ACTTACCTTG	401
	CAGGTAAAGA	ATCGACCTTT	AGGTGCAGTT	CATTAAAAAC	351
GATAAAGTCA	GATTACAGGC	GCAAGCTGAT	GCTAAAGGCG	AAATCAGCAA	301
TTTCCGCTCA	GGCGGTGTAA	AGCGGAAATT	AAGAGGGTGA	CTTTCCGCCA	251
CGAAACCAAG	TGCCACTATT	ATGTCCGTGC	GGTAACATTA	TGCCAAAGGC	151
GCGATATTTT	GTCAATCTGG	AAATGAAGCG	CCGCGCCTGA	TACAGCATTG	101
AACCATTACT	TAATAAACCC	ATCAGCGATA	AAAAATCACC	TCGCAGGGCA	.051
ATTHCHTAC	TGGTGGCAGC	TTAGCGTAAA	GAGGGTGTGA	AGTGAAAAAC	.001
_	GTAAATCTTA	AGACGGCAGT	CTGTCGGTAA	GGTTTAATTA	951
		GATAAAGCGC	GCAAACCAAA	TCACCTTCGA	901
		TTTCTAACGA	ACGCTAGACA	TACGGCTTCT	851
CTAATGGCTT	ATTATTAACA	TAAAGACGCA	TCACAATAGG	CCAAATGGTA	801

#### 'IG. 6C.

AAAGGAAGCA TCAAAAAGGT GACTGCAATT TTTGAAGGGA ACCTAAAAAT ATTTAACCTC GACTCCAGAG AAACGGTATA CAAGAGTCAA	CCTCAGGCAA TCAAAAGGAAGGACCCTCAGGCAGCGGCAGCGGCAGCGGCAGCGGACTCTTTGAAGGGAATTAATT	GGGACTATTA TCTAAACGGC AATACGCTAT GTGAACATCT CAAAGGACGC CAAAGGACGC CTTACCCAGC		ACCAAGTCAT TACAGGTCAA TTTAGATTTA ATAATGTCTC CACCACTAAA AGAACCAATA CTTTAAATAT TTCAGGGAAA GAAAGTGGAT ATGATAAATT GAAAGTGGAT ATGATAAATT GAAGCGATAG TGCAGGCACA TCATTCAACA AAGACACTAC	2101 2151 2201 2251 2301 2401 2451
ACCTAAAAAT		GTGAACATCT		GAAAGTGGAT	2301
TTTGAAGGGA		AATACGCTAT	THO DOCON NO	CTTTAAATAT	2251
GACTGCAATT		TCTAAACGGC	AIAAI'G'I'C'I'C	CACCACTAAA	2201
TCAAAAAGGT		GGGACTATTA		TTTAGATTTA	2151
AAAGGAAGCA	CGCCTTTGAG	AACAAGATAT			2101
GGGCGCAAGG	AICICACI'CG			TAACATAAAC	2051
AACAAT"I"I'AC		ТСАТАААА	GGGTTGATGT	TCAGGCGGCT	2001
GAGALIAACA		GATACCAGAG	CACCGGTGAT	ACGATATTAC	1951
		GTCGGAGCGG	r TGGAGTGAGG	CTTAACTCTT	1901
	ATTAATTAAT	CAATAGCTCC	C GCATCTATGT	GCT'AA'I'CAAC	) )
TAACHACAAT	GTACCTTTGT	CTAAAAAAAG	A'I'A'I'A	)	, ,
ACATTAACAA	AGAAAAGACA				1801
			C AGCACCCCAA	GAATAGTGCC	1751
			C GCAGCAATAC	ACAGCAGGAC	1701
		GACCCGGATA	A GTGGTTGTTA	ACGCCAAAGA	1651

## FIG. 6D.

#### FIG. 6E

*** * * ) * ** <b>!</b> · ) )			-		
CGAAAGTTTA	TTAAAGCAAC	ACAGTAGAGG	TTCTGGTGGC	GCGGTACGAT	1051
GGCGATATCG .	AAGTCAATCA	TAACCACTTC	ACCGAGAGTG	AATTAAAGGA	1001
CAGGCTCTAC	ACCACTTTGG	CGGTGCATTA	CTGCAAATAG	GTTACTGTTA	3951
GGGCAACACC	GCAATATTTC	CTTGCTGTAA	CGAGGGCGCT	TTACTGCAAC	3901
TCTGTAACAC	CAGCTCTGGC	GAATTGAGTC	ATCCTAGGTG	AACAGGTAGT	3851
TAACCGCTCA	AACGTGGAGA	AACCACTGGT	CCATTAACGC	ACAGGTACAA	3801
TACCACTAAA	GTGGAGAAAT	TCTGCGACAA	AGTGAGCATC	CTCACAAAGC	3751
AATATTACTT	AGTAAACAAC	AAAATGTAAC	ATCGATGCAA	CGGCTTAACT	3701
ACAATAATGC	GATAGCAGTG	CAACACTGAA	GTAGTAATAA	GAAACATCCG	3651
CAGCAAAGTG	TGACACTACA	GGTCACAAGG	CTCTGCTGAC	ATTCAAAAAT	3601
	·	CCAAAAAAGT	GGTACTAATG	TAGTGCTGAT	3551
_	TTAACTATTG	TGGTAGTGAT	CAGCTAAAGA	GCAGAGATTA	3501
		CAAGACCTAA	GAAATTAACG	CCAAAGAATT	3451
	TGCCAATCTA	CGACAAACAA	GATTCAGACG	GGAGAATTCC	3401
		ACAGATAACA	. ATATTACCAA	GACAAAATCA	3351
GAТТТСТТСТ	GTAATCTCAC	CAAAAAGAAG	CGATGTCTCG	AAATTGGCGG	3301

## FIG.6F.

4101	ACCACTCAAT	ACCACTCAAT CCAATTCAAA AATTAAAGCA	AATTAAAGCA	ACAACAGGCG AGGCTAACGT	AGGCTAACGT	
	AACAAGTGCA	ACAGGTACAA	TTGGTGGTAC	GATTTCCGGT	AATACGGTAA	
	ATGTTACGGC	AAACGCTGGC	GATTTAACAG	TTGGGAATGG	CGCAGAAATT	
	AATGCGACAG	AAGGAGCTGC	AACCTTAACT	ACATCATCGG	GCAAATTAAC	
	TACCGAAGCT	AGTTCACACA	TTACTTCAGC	CAAGGGTCAG	GTAAATCTTT	
	CAGCTCAGGA	TGGTAGCGTT	GCAGGAAGTA	TTAATGCCGC	CAATGTGACA	
	CTAAATACTA	CAGGCACTTT	AACTACCGTG	AAGGGTTCAA	ACATTAATGC &	_
	AACCAGCGGT	ACCTTGGTTA	TTAACGCAAA	AGACGCTGAG	CTAAATGGCG 8/9	
	CAGCATTGGG	TAACCACACA	GTGGTAAATG	CAACCAACGC	AAATGGCTCC	
	GGCAGCGTAA	TCGCGACAAC	CTCAAGCAGA	GTGAACATCA	CTGGGGATTT	
	AATCACAATA	AATGGATTAA	ATATCATTTC	AAAAACGGT	ATAAACACCG	
	ТАСТСТТААА	AGGCGTTAAA	ATTGATGTGA	AATACATTCA	ACCGGGTATA	
	GCAAGCGTAG	ATGAAGTAAT	TGAAGCGAAA	CGCATCCTTG	AGAAGGTAAA	
	AGATTTATCT	GATGAAGAAA	GAGAAGCGTT	AGCTAAACTT	GGCGTAAGTG	
	CTGTACGTTT	TATTGAGCCA	AATAATACAA	TTACAGTCGA	TACACAAAAT	
	GAATTTGCAA	CCAGACCATT	AAGTCGAATA	GTGATTTCTG	AAGGCAGGGC.	
	GTGTTTCTCA	AACAGTGATG	GCGCGACGGT	GCGCGACGGT GTGCGTTAAT ATCGCTGATA	ATCGCTGATA	

#### 7IG. 6G

4951	ACGGGCGGTA	GCGGTCAGTA	ATTGACAAGG	TAGATTTCAT	CCTGCAATGA
5001	AGTCATTTTA	TTTTCGTATT	ATTTACTGTG	TGGGTTAAAG	TTCAGTACGG
5051	GCTTTACCCA	TCTTGTAAAA	AATTACGGAG	AATACAATAA	AGTATTTTA
5101	ACAGGTTATT	ATTATGAAAA	ATATAAAAG	CAGATTAAAA	CTCAGTGCAA
5151	TATCAGTATT	GCTTGGCCTG	GCTTCTTCAT	CATTGTATGC	AGAAGAAGCG
5201	TTTTTAGTAA	AAGGCTTTCA	GTTATCTGGT	GCACTTGAAA	CTTTAAGTGA
5251	AGACGCCCAA	CTGTCTGTAG	CAAAATCTTT	ATCTAAATAC	CAAGGCTCGC
5301	AAACTTTAAC	AAACCTAAAA	ACAGCACAGC	TTGAATTACA	GGCTGTGCTA
5351	GATAAGATTG	AGCCAAATAA	GTTTGATGTG	ATATTGCCAC	AACAAACCAT
5401	TACGGATGGC	AATATTATGT	TTGAGCTAGT	CTCGAAATCA	GCCGCAGAAA
5451	GCCAAGTTTT	TTATAAGGCG	AGCCAGGGTT	ATAGTGAAGA	AAATATCGCT
5501	CGTAGCCTGC	CATCTTTGAA	ACAAGGAAAA	GTGTATGAAG	ATGGTCGTCA
5551	GTGGTTCGAT	TTGCGTGAAT	TCAATATGGC	AAAAGAAAAT	CCACTTAAAG
5601	TCACTCGCGT	GCATTACGAG	TTAAACCCTA	AAAACAAAAC	CTCTGATTTG
5651	GTAGTTGCAG	GTTTTTCGCC	TTTTGGCAAA	ACGCGTAGCT	TTGTTTCCTA
5701	TGATAATTTC	GGCGCAAGGG	AGTTTAACTA	TCAACGTGTA	AGTCTAGGTT

## FIG. 6H.

CTGCGGGTTT	ACGGTATCCT	AGATATGCAC	CTTACGGCGA	AATGCTAAAA	5551
TAATAGCGAA .	AGTTCCGTTA	GATGCAGGTC	TGCGTTTTAT	TCAGCCCTTA	5501
CGCTTTCAAA	AAAATACACC	TAAGTATGCC	CGTAATGAAT	TCTTGTATGG	5451
GTGAGCGCGG	GGTGCAAGTG	TAAATACGGC	TCAGAGGCTT	ACTTATGGCG	5401
TGTAACAGGT	ATTTATTCTC	AGTAGCATAG	ACAAGATATA	AGTTTACTCT	5351
TTATCGGGTC	TAGCAGTCAA	GTTGGCATTT	TTTGCTCAAG	GAGTCAAGAG	5301
GTTTAGGGTT	AGCACAGCCA	CTATCACATT	TTAATCGCAG	GGCGAAACAT	5251
GGAGCGCATT	CTTTTGGAAT	TTACCAGGCT	CGCGAGTAAA	ATCATTATTA	5201
GATTTAACTC	CTTTAATATT	CTAAAACAAT	CAATTTACCC	TGGACATATC	5151
CAGGCATTGA	GGCGTAAGTG	TGCAGTATCA	AGAAAAATT	GGTGCAACGA	6101
AAACACCCTG	CATCCGAGTT	ATTAATCAAA	CTACCGCCAT	TAGGCTACAA	6051
AAAATTAATT	AGACCAGTTT	TTGGAATGGA	ACATTTAACC	TTATCTCCCG	6001
AATGGAGTTA	GCGAATCTGA	ATCTATCTCT	CAAAAGGTCA	CGTAAATTAT	5951
TGCGATTAAT	GCTTACCAAG	GATATCGACG	TGATTCTAAT	TGAGTTATGC	5901
TATACCAGCA	CTTAAGTCTT	AACACCAATC	TTTTATGATA	TACTTATCCG	5851
GCATAGGATA	TATGCGGTAG	ATCAAAATCT	TAAAAGCACC	TTGACCAATG	5801
TCTAAACGCA	ATGTATTAAA	CAATTTGACC GGACATGATG		TTGTAAATGC	5751

# FIG. 6I.

						2	29/82	2							
GCTTTTGTTG		TCAGTTTCTA	GTTTATAACT	GTTTTCATCC	AAACCAAGCA		ATGCTAAAAA	CAATACAAGG	ACAAAATACG	TGCAAATACT	CATGTCGCCA	TTTGGAAAAA	ACGCACCTGC	GCCACTCGTC	GGCAATTTCC
AAGCTTAGAT	ATTTGAATGG	AGATTAACAT	CCGCCTACCA	ACGCAACCCT			CAAACCAAGC AAACCAAGCA	CATACCATGG	AAAGTGTTCC	CAAACTTCCC	CCTGGAACAA	TAATGGCGAT	ATTGAATTTG	AATTCATTTT	CTCTTTTCCG ACCCCGAATT GGCAATTTCC
CACAAAACTT	AATAGTGACA	CTTCTGGGGT	GGTAAGCGTT	TACAGTCTAT	GCAAACCAAG	CCAAGCAAAC	CAAACCAAGC	AAACATACTC	GAAAATTTAC	AAACAACGAC	ACCTATTACG	TGCCGCGAAT	CGTTCACGAT	AAAAACTACT	CTCTTTTCCG
ACCTCTCCTA	TGCAAATGCC	CACCTACAAC	TTAATCAACT	CCCGCCAATT	TTATATATCA AACAAACTAA	ACCAAGCAAA	GCAAACCAAG	TGATAAACTA AAACATACTC	TATGACAAAA	TTGTAGAATC	CCCAAACCCA	TGAGCTTGCT	ATTTGGAGG	TATCTACCCG	AATTACAACA
AGGCATTAAA	CTCGTCGCTT	CGCACAAGCT	ACCCTGAAAT	ATATGCTTTA	ТТАТАТАТСА	AACCAAGCAA	AGCAAACCAA	ACAATTTATA	GATTTAATAA	ACCGCTTCAC	TAAACAACCA	AAAAAGATTA	ATGGACGCTA	TCAGCTGGCA	TCGCTAATGC
6601	6651	6701	6751	6801	6851	6901	6951	7001	7051	7101	7151	7201	7251	7301	7351

#### FIG. 6J

GATTAGCCTG CAACGCTGGT TGACGCTGAT	TTAACGCAGA CCATATTCTC AATAAATATA	GGTGGCTTTC ATTTAGCAAC AGACAACTCT	TTTTTACTTA CCCGAATCCA ATGTCAATAT	CAGGGAATCA ACAACTTTGT GCTTCATTGT	CGTTTTATTG GTACTGCATC TGCGTTTCAT	GTGGTTTCCT AAAAACTCG CCGAAATTGC 🐱	CAAATATCCT TCATGATGTA TATATGCACT 😿	AACAAGCACG ATGTTAAGCG TCCATTAAAC	CCTCACGCAA GGATGGCAAG ACCGCTACCT	ACGGCAAACC TGTGATGATG GTACTGCTTG	TCGATTTATC GCACGCATTC AACTTCAATG	CTATTTAGTC GGCTTAGGCC ATGAGGGCGT	TGTTTGACGA GTTCTTTGAA ATCAGTAGCA	TTTTTTATCC GTAAACAGTG CGAAACTTTC .	
CATTAAAGAT GATTA	TCCCCCTACG TTAAC	AGATTCCGAA GGTGG	AATTCTGTAT TTTTT	GCGTTATGGG CAGGG,	GCAGTCTTCA CGTTT	TGGTTTTACA GTGGT	GAATTGCCTG CAAAT,	TTTAGCAAAA AACAA	GCAAGCATAT CCTCA	GGTAAAAGG ACGGC	TTCGGGACAT TCGAT	GAGAAAATT CTATT	GGTCGAGAAG TGTTT	GGAGAGACTG TTTTT	
GAAGAAGGGG	TTTGCCTCT 1	ATATCAACCC A	TCTATTGCTA A	GAGTTTAGAT C	GTTTTGCGTT (	AAAAGAGCGG 1	TAATTTAGAT G	GCAGTTATGA 1	GAACTTGTCC 0	TTACACCTTA 0	AACATTTTAA 1	ATTGCTGCTC 0	TGATAACATA G	ATAATATAAT C	
7401	7451	7501	7551	7601	7651	7701	7751	7801	7851	7901	7951	8001	8051	8101	

#### IG. 6K

GCC.11.CCC.1.C	CGTAGAAGAT	TACGCTTACC	CAAAAGTGG	TATTGCCGCT	ATCAG	ATCAA	31/2 TAGGT		TACTA	AAACG	TTAGG	ATGTG	STTAC	JGTCC	CACT	Caaa
		TACGC		TATTG	AAGAAATCAG	GGACAATCAA	CTATTTAGGT	ATCTGGCAAT	GGTAATACTA	ATGCAAAACG	AACGCTTAGG	ATTGAATGTG	CCGTCGTTAC	ACCCTCGTCC	CGGAAGCACT	AATTT
	ATTATGTCAT	GAAACCCTTT	ACTCGCCCCA	TCAATATCGG	CTAACATTGC	TTTCGCACTT	TTATCGAAAG	TATCACGATT	GTTTCCTTTC	TAGTTGGTGT	GGTCTGTTTA	AGAAACATAT	GCCTTGAACT	TTTACAGGCG	TGAATGGAAG	AAGTGCGGTT AATTTTCAAA
	GAATTTATTG	TTGTTTAGC	TACCATCTGC	CCTGAAGTAG	TGAATTTTTG	TACATTTTCA	GTCAAATGGT	CCACGCACCT	TACTAAATCC	ACATTAGGTT	TATTGATGAA	CCGACACACG	CATCAAGAAC	ACAAAAGCTT	AGAAAACAAA	TTTAAAGTAA
	TACGCATTCT	GCAGTGAAGA	CTACCTTATG	CAGGGAAAAC	AATTAAACCC	AAAGTCAAAA	ACACCCTTAT	CTGCACATCC	TGCGATATGC	TGATATGGTT	TACATGAACA	TGGCTGATAG	AGCAGAAAAC	ACAACGGCTT	ATACTGCTTA	ATAACGGTTT
	ATCCTGCCAC	GATTATGTGG	CAAAGATGCC	ATTATGTACT	ACCACAATGA	AGATAAAGCT	CAGGCTTGAC	GACGATGCCA	ATTGCGTGAT	ACGGCATAAT	GGGGATGAAG	ACTACCAGAA	CTTTGCGTCT	ATCATAGAAA	ATTGGGCAAA	TGAGTAAAAA
	8251	8301	8351	8401	8451	850.1	8551	8601	8651	8701	8751	8801	8851	8901	3951	9001

#### 32/82

## FIG. 6L.

TCCCGCGCGC TGACAGTTTA TCTCTTTCTT AAAATACCCA TAAAATTGTG GCGTTTTAAA AACCTCTCAA AAATCAACCG CACTTTTATC TTTATAACGC GCAATAGTTG GGTAATCAAA TTCAATTGTT GATACGGCAA ACTAAAGACG GCGCGTTCTT CGGCAGTCAT C 9051 9101 9151 9201

### FIG. 7A.

ATTTTGGATT GTTGAAATTC AACTAACCAA AAAGTGCGGT TAATTGTTCA CTTTTTCGGT ACAACTTTAC GCGAATACGT CATAATAGGT TTTGCAAGAT TTTTAATTCA TAAATATACA ACAATTACAA AGTATAAATC ATCTTTCATC GATGAACCGA TCTTTCATCT GAACGCAAAT TATGAACAAG CAATCCACCA AAGAACGAGG GCTCTTCTTG ATGACAAACA GTTGGCGTTT TCGGGATAAT AATCATAAAT AAATCGCCAA TTAAAAAAAT ATCAACTGGT TTCATCTTTC TCATCTTTCA CACATGAAAT GGAGCTGAAC TAGGAGAAAA GGTTGTAGTG TTGGGCATTG GACGACTATG GTAAGTTCTT TTTGTTTAGC AAGAAAATGA GATGTTCTAA ATTTGTGGCG TCCCACTCAA GATTTTTGTG TGCAAATATT CAACTAACCT TCTTTCATCT CTTTCATCTT TTTCATCTTT GAATGAAGAG GGAGAAAATA GCTCTCTTAA TTATATTCTG AAGCGTTAAT GCATAATATT AATAAATTTT TTCAATACCT AATGGTATAA AATAAATCAA GCAGTCTATA GGAGGGGCAA CATCTTTCAT ATCTTTCATC TTTAATTGTT CGCCACTTCA AAAGGATAAA TAAAATCTGT CGTTGGTTTT GTTGCCCAAA ATTTCTTGTA AGATAATAAA CGCCATATAA TAATAGTAAA AATCCCATTT ATTGTGGCAA GGGAAGGGAG CACCTTTTTT TTTCATCTTT TTCATCTTTC GATAAAGTAA Н 51 101 151 201 251 301 351 401 451 551 501 601 651 701 751

#### FIG. 7B

801	ATATATCGTC	TCAAATTCAG	CAAACGCCTG	AATGCTTTGG	TTGCTGTGTC
851	TGAATTGGCA	CGGGGTTGTG	ACCATTCCAC	AGAAAAAGGC	AGCGAAAAAC
901	CTGCTCGCAT	GAAAGTGCGT	CACTTAGCGT	TAAAGCCACT	TTCCGCTATG
951	TTACTATCTT	TAGGTGTAAC	ATCTATTCCA	CAATCTGTTT	TAGCAAGCGG
1001	CAATTTAACA	TCGACCAAAA	TGAÀATGGTG	CAGTTTTAC	AAGAAAACAA
1051	GTAATAAAAC	CATTATCCGC	AACAGTGTTG	ACGCTATCAT	TAATTGGAAA
1101	CAATTTAACA	TCGACCAAAA	TGAAATGGTG	CAGTTTTAC	AAGAAAACAA
1151	CAACTCCGCC	GTATTCAACC	GTGTTACATC	TAACCAAATC	TCCCAATTAA
1201	AAGGGATTTT	AGATTCTAAC	GGACAAGTCT	TTTTÄATCAA	CCCAAATGGT
1251	ATCACAATAG	GTAAAGACGC	AATTATTAAC	ACTAATGGCT	TTACGGCTTC
1301	TACGCTAGAC	ATTTCTAACG	AAAACATCAA	GGCGCGTAAT	TTCACCTTCG
1351	AGCAAACCAA	AGATAAAGCG	CTCGCTGAAA	TTGTGAATCA	CGGTTTAATT
1401	ACTGTCGGTA	AAGACGGCAG	TGTAAATCTT	ATTGGTGGCA	AAGTGAAAAA
1451	CGAGGGTGTG	ATTAGCGTAA	ATGGTGGCAG	CATTTCTTTA	CTCGCAGGGC
1501	AAAAAATCAC	CATCAGCGAT	ATAATAAACC	CAACCATTAC	TTACAGCATT
1551	GCCGCGCCTG	AAAATGAAGC	GGTCAATCTG	GGCGATATTT	TTGCCAAAGG

#### FIG.7C

35/82																
GGTAAACTTT	TCTTTCCGCC	-				AAAGAAAAAG			AAAACAAAAG		CCGGTGAAGC	ACCAATACAA	AACGGCATCA			CGGATGGGTT
TCGAAACCAA	GCAATATTGT	ATTTCCGCTC	CGATAAAGTC			TGTATCAGGC	CGTTAATTGA	ACCGGTGGTT	TGCAATTGTT	TTGAAGCCGA	CCAACAGGCA	AACAACGCTA	CAATGAATAT	ATCGGAAGCA	AGGCGTTCAG	TTTATTCTGG
CTGCCACTAT	GATAAAAGCG	TGGCGGTGTA	TGATTACAGG	TCAGGTAAAG	AGGTAAAAAC	CAACCATCAA	GGCGATATTG	TATCGCTAAA	TTGACAGCAA	GATGTAACAA	TGATGAATTC	GCGAACTCAA	AACGCCTGGA	CTCAATCAAC	AGĊGTGGCGG	AATTTAACCA
AATGTCCGTG	TGTAAGCAAA	AAGCGGAAAT	GGCAAGCTGA	TATCGACCTT	AGCGCGGCGA	GAAAAAGGCT	TATTGTGTGG	GTAGTGGTGA	TATTTATCCA	AGACCCTGAT	CCGGTATAAA	AAAAAAATA	TTATCTGAAA	CCGTTAATAG	AGTAAAGGTC	TAAAGGCGGA AATTTAACCA TTTATTCTGG
CGGTAACATT	CTGCTGATTC	AAAGAGGGTG	AGCTAAAGGC	CAGGTGCAGT	GGCGGTGACG	AACCTCTTTA	GCGGACGCGC	AACGCTCAAG	ATCGGGGCAT	AGTGGTTGCT	CGCAATAATA	AAGCGACCCT	CTATTTCAAA	AGAAAACTTA	AATTCTCCAT	ATATTACTTC
1601	1651	1701	1751	1801	1851	1901	1951	2001	2051	2101	2151	2201	2251	2301	2351	2401

## FIG.7D.

CCGC		נאנ) א	6.000 6.000		ST TO		5/82 5 80 5 00 5 00 5 00 5 00 5 00 5 00 5 00	2000	\TT\	TTC	GAT	AAT	90 90	TCA	ACG
ATATTACCGC			AAGGTOTGAA	GGCACAATTA	GAACAC	CTCTTAATCT	AGCAATAGCA	TTTTAACGGC	AAGTTAATTT	TTACCAATTC	TTTTTTGAT	TGAGTGAAAT	GTTCGCGGCG	AACCAATTCA	GGTACGCACG
GGTTTTTAA	CAAAGCACGC	CCATTACAGG	GGAACGGGTA	CAATCTTAGT	CTACGAGAAA GAACACCTCG	AACGTCAGTG	ATACATTTCA	CAGGGGTGAA	CAATCTCAAA GAAGGAGCGA	AAGCAAACCT	GGGCTCTGT	GAGTTAAAAA	AAATTCCCAT	CCATAAATGC	TTTTATGACG
GCTTGATCAG	GTGGAAATAA	GGCACTGTAA	ATCTTTAAAC	ATTTAACCCA	ATTAACCAAA	TTCGCACTGG	CCTTTATTAA	AGAAGCTCTG	CAATCTCAAA	ACATGAACAC	GCCACTGGTG	CAGAGGGGCT	ATTTACCTT	AAAGACTTAA	
AAAATATTAC	GCTTTTGAAG	TGTCGCCCAG	CTAACAACGT	TCAGTGAATA	GAATATAACA	CCÁGCCATGA	GCAAATTTTA	AACACAGTAT	ACATGTCATT	CCAAACGAGA	CAATATCACA	ACCATTCTGG	AACGGCGCTA	TAAAATCAAC	TCAGACAGAC GAAAGATGAT
GATGTTCATA	CGCTTCCGTA	ATGCTAAAAT	GATTTCAGGG	TATCATTTCA	ACATATCTGG	TATTGGCAAA	AGAGACAGGC	AAGGCTTAAC	GTAAATGGCA	CAAATTAAAA	GGTTTTAGC	ATATATGCCA	TAATATCTCT	ATGACGCTTT	AATTTCAGCC
742T	2501	2551	2601	2651	2701	2751	2801	2851	2901	2951	3001	3051	3101	3151	3201

# FIG.7E.

37/82 GGTAATGTCA TATTACTATC CTAATCAGCA GTTAATGGGA TCTCACTATT TAAATATCAC ACACAAGGAG CATTACCACT TAATCAACAA GAAATCCAAA TTCTTCCGAT TTGATGGAGA ATTAAAACCA CAATAAAGCA ACAGTAATGA AATGTTAAAG CATTCTGGGC AATAACGCCC CAGCTTGCTC TTACGGGGAA TTAAAGGCAA AGGAAAGACT AGAGATACCC AATTAATA GTGATTTAAA GGCGGAGATA TAATGATGCT ACCTCACGAT AAAAAGGGTA CAACCTAACT AACTTTTAAC ACTATTGGCA TTTCAGGTTT CAATGCCATC AATTCAACCT ACAACATATC GCTAGAAGCC AGCAGCAGCA TAAAACTTGG AATGCAGATA GCACTGCCGA TATCTCGCAA AAAGAAGGCA ACCAATGATG AAGCATCATC CAGACAGTAA GATAACAATC CAAGTAATGC GACCTAAGTA TAGAGATTTA CCAAAACAGT ACAAAACTCA AACTGGCGAA CCACTTTTAA CAAATGTTAC GATAGAGTTA TGGCAATGTT CGGCAATTTT ACCAATAATG GCAACCAAAG TTAAATATTA TCACCAAACA TCAGATGCGA ATTGACAGAA CCAAAGATGG GGTGCCGAAG GAGAAAGCAG CCCTTGGTGG AAACATAAGG GTTTAAGTTT TCAGAAAGCG TGGTAAAACT CACGCTAAAC AAAAGGAAGC TTGGCGGCAA AAAATTAATA GGACTCTAGT AAGAATTGAA GAGATTACAG CGGTAACAGC 3251 3301 3351 3401 3451 501 3651 551 3601 3701 3751 3801 851 3901 3951 4001

# FIG.7F

							38	8/82							•		
CHO K K K K K C TO A C F			T.I.DWI.I.MO		CAACCAAAAC	GITAGUGUGA			GIIGGGAATG	CGCAACAGGG	CTAAGGGTCA	ATTAATGCTG	GGCAGGCTCC	AAGATGCTAA			gaaagargg
TGACACTAAA								41990001111 4044FFF480		CANCITAAC	ATCACTTCAA	CGCAGGAAGC					
GGTCACAATG	CGGACGTGAA		ACCGCGTCGG			TCCGGCTCAA	AACAGGTACA				CGGTTCTAGC	ATGGTAGCAT	ACAGGCACCT	CACCTTGGTT	GTGATAGTAC		AAATGGGTTA
ATTCAAAAAT CTCTGCTGAC GGTCACAATG	GCAGCAATGG	ATTACTGCAA	AGTAAATATC				TAACAAGTGC	AATGTTACGG	TAATGCGACA	00 k k 0 E 0 k E 0	CIACIGAAGO	TTGGCTCAGA	ATTAAATACT	CAACCAGCGG	GATGCATCAG	TGGTAGTGTG	TAAACACAGT
ATTCAAAAAT	AAAACATCTA	CGGCTTAACT	CTCTCAAAAC	GGCTCGACCA	AGGTGATATC	CTGGTGATTT	GAGGCTAATG	TAATACGGTA	GCGCAGAAAT	AATACCHUCA	UD T T O O T T T T T	GGTAGACCTC	CTAATGTGAC	GATATTAAAG	GCTAAATGGT	ACTGGGGATT	ACTGGGGGATT
4051	4101.	4151	4201	4251	4301	4351	4401	4451	4501	4551	1 .	4601	4651	4701	4751	4801	4851

# FIG.7G.

						3	39/82	<u>)</u>							
АААТАТАТС					TATGTACAA	GTAGATTTCA	GTGGGTTAAA	GAATACAATA	GCAGATTAAA	TCATTGTATG	TGCACTTGAA	ТАТСТАААТА	CTTGAATTAC	GATATTGCCG	TCTCGAAATC
AATTGAGGTG		·				AATTGACAAG	TATTTACTGT			GGCTTCTTCA	AGTTATCTGG	GCAAAATCTT	AACAGCACAG	AATTTGATGT	TTTGAGCTAG
GAGGCAAGGA	GAAGAAGTAA	TGATGAAGAA	TTGTTGAGCC	ACCAGACCGT	AAGTGGTAAT	CGTAGTCAGT	ATTTTCGTAT	ATCTTGTAAA	TATTATGAAA	TGCTTGGCCT	AAAGGCTTTC	ACTGTCTGTA	CAAACCTAAA	GAGCCAAATA	CAATATCATG
GTGCGCTTAA	AGCAAGTGTA	AAGATTTATC	GCTGTACGTT	TGAATTTACA	CGTGTTTCTC	GATGGACAGC	AAGTCATTTT	GGCTTTACCC	AACAGGTTAT	ATATCAGTAT	GTTTTTAGTA	AAGACGCCCA	CAAACTTTAA	AGATAAGATT	TTACGGATGG CAATATCATG
TAGAAACACT	AGCCAGGTGT	GAAAAAGTAA	TGGTGTAAGT	ATACACAAAA	GAAGGTAAGG	TGTTGCTGAC	TCCTGCAATG	GTTCAGTACG	AAGTATTTT	ACTCAGTGCA	CAGAAGAAGC	ACTTTAAGTG	CCAAGGCTCG	AGGCTGTGCT	CAACAAACCA
4901	4951	5001	5051	5101	5151	5201	5251	5301	5351	5401	5451	5501	5551	5601	5651

# FIG.7H.

5701	AGCCGCAGAA	. AGCCAAGTTT	TTTATAAGGC	GAGCCAGGGT	TATAGTGAAG	
5751	AAAATATCGC	TCGTAGCCTG	CCATCTTTGA			
5801	GATGGTCGTC	AGTGGTTCGA			44 4 4 4 4 4 4 A A A A A A A A A A A A	
5851	CCCGCTTAAG	GTTACCCGTG	_			
5901	CCTCTAATTT	GATAATTGCG	_		A A CCCCE A CALA	
5951	TTTATTTCTT	ATGATAATTT	CGGCGCGAGA	GAGTTTA A CT	TACCC LACC	
6001	AAGCTTGGGT	TTTGTTAATG	CCAATTTAAC	TGGTCATGAT		
6151	TTATACCAGT	ATGAGTTATG	CTGATTCTAA	TGATATCGAC		40
6201	GTGCGATTAA	TCGTAAATTA	TCAAAAGGTC	AATCTATCTC		/82
6251	AAATGGAGTT	ATTATCTCCC	AACATTTAAC	ひて見べしひむ氏し		
6301	ТААААТТААТ				AAGACCAA'I"!'	
' ',	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	474175541	ALIACCECCA	TATTAATCAA	ACCTCCGCGT	
6351	TAAATCGCTT	GGGTGAAACG	AAGAAAAAT	TTGCAGTATC	AGGCGTAAGT	
5401	GCAGGCATTG	ATGGACATAT	CCAATTTACC	CCTAAAACAA	ТСТТТААТАТ	
5451	TGATTTAACT	CATCATTATT	ACGCGAGTAA		TCTTTTCGAA	
5501	TGGAGCGCAT	TGGCGAAACA	TTTAATCGCA		TAGCACACC	
5551	AGTTTAGGGT	TGAGTCAAGA	GTTTGCTCAA		TTAGCAGTCA	
5601	ATTATCAGGT	CAATTTACTC	TACAAGATAT		GATTTATTCT	

## FIG.7I

6651	CTGTAACAGG	TACTTATGGC	GTCAGAGGCT	TTAAATACGG	CGGTGCAAGT
6701	GGTGAGCGCG	GTCTTGTATG	GCGTAATGAA	TTAAGTATGC	CAAAATACAC
6751	CCGCTTCCAA	ATCAGCCCTT	ATGCGTTTTA	TGATGCAGGT	CAGTTCCGTT
6801	ATAATAGCGA	AAATGCTAAA	ACTTACGGCG	AAGATATGCA	CACGGTATCC
6851	TCTGCGGGTT	TAGGCATTAA	AACCTCTCCT	ACACAAAACT	TAAGCCTAGA
6901	TGCTTTTGTT	GCTCGTCGCT	TTGCAAATGC	CAATAGTGAC	AATTTGAATG
6951	GCAACAAAA	ACGCACAAGC	TCACCTACAA	CCTTCTGGGG	GAGATTAACA
7001	TTCAGTTTCT	AACCCTGAAA	TTTAATCAAC	TGGTAAGCGT	TCCGCCTACC
7051	AGTTTATAAC	TATATGCTTT	ACCCGCCAAT	TTACAGTCTA	TAGGCAACCC
7101	TGTTTTACC	CTTATATATC	AAATAAACAA	GCTAAGCTGA	GCTAAGCAAA
7151	CCAAGCAAAC	TCAAGCAAGC	CAAGTAATAC	TAAAAAAACA	ATTTATATGA
7201	TAAACTAAAG	TATACTCCAT	GCCATGGCGA	TACAAGGGAT	TTAATAATAT
7251	GACAAAAGAA	AATTTGCAAA	ACGCTCCTCA	AGATGCGACC	GCTTTACTTG
7301	CGGAATTAAG	CAACAATCAA	ACTCCCCTGC	GAATATTTAA	ACAACCACGC
7351	AAGCCCAGCC	TATTACGCTT	GGAACAACAT	ATCGCAAAAA	AAGATTATGA
7401	GTTTGCTTGT	CGTGAATTAA	TGGTGATTCT	GGAAAAAATG	GACGCTAATT

## FIG.7J

7451	TTGGAGGCGT	TCACGATATT	GAATTTGACG	CACCCGCTCA	GCTGGCATAT
7501	CTACCCGAAA	AATTACTAAT	TTATTTGCC	ACTCGTCTCG	CTAATGCAAT
7551	TACAACACTC	TTTTCCGACC	CCGAATTGGC	AATTTCTGAA	GAAGGGGCGT
7601	TAAAGATGAT	TAGCCTGCAA	CGCTGGTTGA	CGCTGATTTT	TGCCTCTTCC
7651	CCCTACGTTA	ACGCAGACCA	TATTCTCAAT	AAATATAATA	TCAACCCAGA
7701	TTCCGAAGGT	GGCTTTCATT	TAGCAACAGA	CAACTCTTCT	ATTGCTAAAT
7751	TCTGTATTTT	TTACTTACCC	GAATCCAATG	TCAATATGAG	TTTAGATGCG
7801	TTATGGGCAG	GGAATCAACA	ACTTTGTGCT	TCATTGTGTT	TTGCGTTGCA
7851	GTCTTCACGT	TTTATTGGTA	CCGCATCTGC	GTTTCATAAA	AGAGCGGTGG
7901	TTTTACAGTG	GTTTCCTAAA	AAACTCGCCG	AAATTGCTAA	TTTAGATGAA
7951	TTGCCTGCAA	ATATCCTTCA	TGATGTATAT	ATGCACTGCA	GTTATGATTT
8001	AGCAAAAAC	AAGCACGATG	TTAAGCGTCC	ATTAAACGAA	CTTGTCCGCA
8051	AGCATATCCT	CACGCAAGGA	TGGCAAGACC	GCTACCTTTA	CACCTTAGGT
8101	AAAAAGGACG	GCAAACĊTGT	GATGATGGTA	CTGCTTGAAC	ATTTTAATTC
8151	GGGACATTCG,	ATTTATCGTA	CACATTCAAC	TTCAATGATT	GCTGCTCGAG
8201	AAAAATTCTA	TTTAGTCGGC	TTAGGCCATG	AGGCGTTGA	TAAAATAGGT

### FIG.7K

8251	CGAGAAGTGT	TTGACGAGTT	CTTTGAAATC	AGTAGCAATA	к O O T A A T A A T A
8301	GAGACTGTTT	TTTATCCGTA			
8351	TCTATATGCC	·			
ν 107α					TGTGAGCAAC
ナ サ つ	ACT GGCT TG	CCCCTATTCA	AGCTGTAGCC	CTGGGTCATC	CTGCCACTAC
8451	GCATTCTGAA	. TTTATTGATT	ATGTCATCGT	·	
8501	GTGAAGATTG	TTTCAGCGAA	ACCCTTTTAC		
8551	CCTTATGTAC	CTTCTGCACT			
8601	GGAAAACCCT	GAAGTAGTCA	ATATCGGTAT		
8651	TAAACCCTGA	ATTTTTTCT			ACAATGAAAT
7	: :		044701100tz	AAATICAGAGA	TAAAGCTAAA
T0/8	GTCAAAATAC	ATTTTCATTT	CGCACTTGGA	CAATCAACAG	GCTTGACACA
8751	CCCTTATGTC	AAATGGTTTA	TCGAAAGCTA	TTTAGGTGAC	
8801	CACATCCCCA	CGCACCTTAT	CACGATTATC	TGGCAAAAAGG	
8851	GATATGCTAC	TAAATCCGTT			
0			) ) ;	DATACT AACG	GCAT'AAT'TGA
T O A G	TATGGT"IACA	TTAGGTTTAG	TTGGTGTATG	CAAAACGGGG	GATGAAGTAC
3951	ATGAACATAT	TGATGAAGGT	CTGTTTAAAC	GCTTAGGACT	ACCAGAAATGG
9001	CTGATAGCCG	ACACACGAGA	AACATATATT		
9051	AGAAAACCAT	CAAGAACGCC	しつしましょ ないまし		
		)	りつつてつじょう	LCGTTACATIC .	ATAGAAAACA

# FIG.7L.

ACGGCTTACA AAAGCTTTTT ACAGGCGACC CTCGTCCATT GGGCAAAATA GTAAAAAATA TTTTAAAAAC TTTTATCTTT ATAACGATCC CGCACGCTGA GCGGAGATTT TTGCACCACA AAGCACTTGA TGCGGTTAAT TTTCAAAGCG TAAAGGCTAA AATCACCAAA GCCTTTCATG ATGGAAGCGG CAGTTTATCA GCCTCCCGCC ATAAAACTCC TACCCACAAA AAA CTGCTTAAGA AAACAAATGA AAAGTAAAAG TCAACCGCAC TGGCAGAAAT ACGGTTTTTT CTCTCAAAAA TAGCCAAAAC AAATCACCAA 9101 9151 9301 9201 9251 9351 9401

# FIG.8A

1 ATGAACAAGA TATATCGTCT CAAATTCAGC AAACGCCTGA ATGCTTTGGT

TOCTGTGTCT GAATTGACAC GGGGTTGTGA CCATTCCACA GAAAAGGCA

101 GTGAAAACC TGTTCGTACG AAAGTACGCC ACTTGGCGTT AAAGCCACTT

TCCGCTATAT TGCTATCTTT GGGCATGGCA TCCATTCCGC AATCTGTTT

AAGTAGACGG CAATAAAACC ACTATCCGTA ATAGCGTCAA TGCTATCATC 201 AGCGAGCGGT TTACAGGGAA TGAGCGTCGT ACACGGTACA GCAACCATGC

301 AATTIGGAAAC, AATTITAACAT TGACCAAAAT GAAATGGTGC AGTTTTTACA

AGAAAGCAGC AACTCTGCCG TTTTCAACCG TGTTACATCT GACCAAATCT

45/82

CCAAATGGTA TCACAATAGG TAAAGACGCA ATTATTAACA CTAATGGCTT 401 CCCAATTAAA AGGGATTTTA GATTCTAACG GACAAGTCTT TTTAATCAAC

501 TACTGCTTCT ACGCTAGACA TITCTAACCA AAACATCAAG GCGCGTAATT

TCACCCTTGA GCAAACCAAG GATAAAGCAC TCGCTGAAAT CGTGAATCAC

601 GGTTTAATTA CCGTTGGTAA AGACGGTAGC GTAAACCTTA TTGGTGGCAA

AGTGAAAAAC GAGGCGTGA TTAGCGTAAA TGGCGGTAGT ATTTICTTTAC

701 TTGCAGGGCA AAAAATCACC ATCAGCGATA TAATAAATCC AACCATCACT

TACAGCATTG CTGCACCTGA AAACGAAGCG ATCAATCTGG GCCATATTTT

## FIG 8B

801 TGCCAAAGGT GGTAACATTA ATGTCCGCGC TGCCACTATT CGCAATAAAG

GTAAACTITIC TGCCGACTCT GTAAGCAAAG ATAAAAGTGG TAACATTGTT

901 CTCTCTGCCA AAGAAGGTGA AGCGGAAATT GGCGGTGTAA TTTCCGCTCA

AAATCAGCAA GCCAAAGGTG GTAAGTTGAT GATTACAGGC GATAAAGTTA

1001 CATTGAAAAC GGGTGCAGTT ATCGACCTTT CGGGTAAAGA AGGGGGAAAA

ACTTATCTTG GCGCTGACGA GCGTGGCGAA GGTAAAAACG GCATTCAATT

1101 AGCAAAGAAA ACCACTTTAG AAAAAGGCTC AACAATTAAT GTGTCAGGTA

AAGAAAAAGG TGGGCGCCT ATTGTATGGG GCGATATTGC GTTAATTGAC

46/82

1201 GGCAATATTA ATGCCCAAGG TAAAGATATC GCTAAAACTG GTGGTTTTGT

GCACACCTCG GGGCATTACT TATCCATTGA TCATAACGCA ATTGTTAAAA

1301 CAAAAGAATG GCTACTAGAC CCAGAGAATG TGACTATTGA AGCTCCTTCC

GCTTCTCGCG TCGAGCTGGG TGCCGATAGG AATTCCCACT CGGCAGAGGT

1401 GATAAAAGTG ACCCTAAAAA AAAATAACAC CTCCTTGACA ACACTAACCA

ATACAACCAT ITCAAATCIT CICAAAAGIG CCCACGIGGI GAACATAACG

1501 GCAAGGAGAA AACTTACCGT TAATAGCTCT ATCAGTATAG AAAGAGGCTC

CCACTTAATT CTCCACAGTG AAGGTCAGGG CGGTCAAGGT GTTCAGATTG

1601 ATAAAGATAT TACTTCTGAA GGCGGAAATT TAACCATTTA TTCTGGCGGA

TGGGTTGATG TTCATAAAAA TATTACGCTT GGTAGCGGCT TTTTAAACAT

1701 CACAACTAAA GAAGGAGATA TCGCCTTCGA AGACAAGTCT GGACGGAACA

ACCTAACCAT TACAGCCCAA GGGACCATCA CCTCAGGTAA TAGTAACGC

1801 TITAGAITTA ACAACGICTC TCTAAACAGC CTTGGCGGAA AGCTGAGCTT

TACTGACAGC AGAGGGACA GAGGTAGAAG AACTAAGGGT AATATCTCAA

AAAGCACCCA AAGTCAGCTG GTTTTACAGA GACAAAGGAC GCACCTACTG

1901 ACAAATITGA CGGAACGITA AACATITCCG GAACTGTAGA TATCTCAATG

47/82

2001 GAACGTAACC ACTITTAAATG ITTACCTCGGG TAGTAAATTT AACCTCTCCA

ITGACAGCAC AGGAAGTGGC TCAACAGGTC CAAGCATACG CAATGCAGAA

2101 TTAAATGGCA TAACATTTAA TAAAGCCACT TTTAATATCG CACAAGGCTC

AACAGCTAAC TTTAGCATCA AGGCATCAAT AATGCCCTTT AACAGTAACG

2201 CTAACTACGC ATTATTTAAT GAAGATATTT CAGTCTCAGG GGGGGGTAGC

CITAATITICA AACITAACGC CICATCIAGC AACATACAAA CCCCTGGCGT

3301 AATTATAAAA TCTCAAAACT TTAATGTCTC AGGAGGGTCA ACTTTAAATC

TCAAGGCTGA AGGTTCAACA GAAACCGCTT TTTCAATAGA AAATCATTTA

### FIG.8D

2401 AACTTAAACG CCACCGGTGG CAATATAACA ATCAGACAAG TCGAGGGTAC

CGATTICACGC GTCAACAAG GTGTCGCAGC CAAAAAAAAC ATAACTTTTA

2501 AAGGGGGTAA TATCACCTTC GGCTCTCAAA AAGCCACAAC AGAAATCAAA

GGCAATGTTA CCATCAATAA AAACACTAAC GCTACTCTTT GTGGTGCGAA

2601 TITIGCCGAA AACAAATCGC CTITAAATAT AGCAGGAAAT GITATIAATA

ATGGCAACCT TACCACTGCC GGCTCCATTA TCAATATAGC CGGAAATCTT

2701 ACTGTTTCAA AAGGCGCTAA CCTTCAAGCT ATAACAAATT ACACTTTTAA

TGTAGCCGGC TCATTTGACA ACAATGGCGC TTCAAACATT TCCATTGCCA

48/82

2801 GAGGAGGGC TAAATTTAAA GATATCAATA ACACCAGTAG CTTAAATATT

ACCACCAACT CTGATACCAC TTACCGCACC ATTATAAAAG GCAATATATC

2901 CAACAAATCA GGTGATTTTGA ATATTATTTGA TAAAAAAAGC GACGCTGAAA

TCCAAATTGG CGGCAATATC TCACAAAAAG AAGGCAATCT CACAATTTCT

3001 TCTGATAAAG TAAATATTAC CAATCAGATA ACAATCAAAG CAGGCGTTGA

AGGGGGGGT TCTGATTCAA GTGAGGCAGA AAATGCTAAC CTAACTATTC

3101 AAACCAAAGA GTTAAAATTG GCAGGAGACC TAAATATTTC AGGCTTTAAT

AAAGCAGAAA TTACAGCTAA AAATGGCAGT GATTTAACTA TTGGCAATGC

## FIG.8E

3201 TAGCCGTGGT AATGCTGATG CTAAAAAGT GACTTTTGAC AAGGTTAAAG

ATTCAAAAAT CTCGACTGAC GGTCACAATG TAACACTAAA TAGCGAAGTG

3301 AAAACGTCTA ATGGTAGTAG CAATGCTGGT AATGATAACA GCACCGGTTT

AACCATTITCC GCAAAGATG TAACGGTAAA CAATAACGTT ACCTCCCACA

3401 AGACAATAAA TATCTCTGCC GCAGCAGGAA ATGTAACAAC CAAAGAAGGC

ACAACTATCA ATGCAACCAC AGGCAGCGTG GAAGTAACTG CTCAAAATGG

3501 TACAATTAAA GGCAACATTA CCTCGCAAAA TGTAACAGTG ACAGCAACAG

AAAATCTTGT TACCACAGAG AATGCTGTCA TTAATGCAAC CAGCGGCACA

3601 GTAAACATTA GTACAAAAC AGGGGATATT AAAGGTGGAA TTGAATCAAC

TICCGGTAAT GTAAATATTA CAGCGAGCGG CAATACACTT AAGGTAAGTA

49/82

3701 ATATCACTGG TCAAGATGTA ACAGTAACAG CGGATGCAGG AGCCTTGACA

ACTACAGCAG GCTCAACCAT TAGTGCGACA ACAGGCAATG CAAATATTAC

3801 AACCAAAACA GGIGATAICA ACGGIAAAGT TGAATCCAGC TCCGGCTCTG

TAACACTIGI IGCAACTGGA GCAACTCITG CIGIAGGIAA TAITITCAGGI

3901 AACACTGTTA CTATTACTGC GGATAGCGGT AAATTAACCT CCACAGTAGG

TICTACAATT AATGGGACTA ATAGTGTAAC CACCTCAAGC CAATCAGGGG

## FIG.8F

4001 ATATTGAAGG TACAATTTCT GGTAATACAG TAAATGTTAC AGCAAGCACT

GGTGATITIAA CTATIGGAAA TAGTGCAAAA GTIGAAGCGA AAAATGGAGC

1101 TGCAACCTTA ACTGCTGAAT CAGGCAAATT AACCACCCAA ACAGGCTCTA

GCATTACCTC AAGCAATGGT CAGACAACTC TTACAGCCAA GGATAGCAGT

TTTAACTACT ACAGGGATT CAAAGATTAA CGCAACCAGT GGTACCTTAA 1201 ATCGCAGGAA ACATTAATGC TGCTAATGTG ACGTTAAATA CCACAGGCAC

301 CAATCAATGC AAAAGATGCC AAATTAGATG GTGCTGCATC AGGTGACCGC

401 AACCTCAAGC AGCGTGAATA TCACCGGGGA TTTAAACACA ATAAATGGGT

TAAATATCAT ITCGGAAAAT GGTAGAAACA CIGTGCGCTT AAGAGGCAAG

501 GAAATTGATG TGAAATATAT CCAACCAGGT GTAGCAAGCG TAGAAGAGGT

AATTGAAGCG AAACGCGTCC TTGAGAAGGT AAAAGATTTA TCTCATCAAG

iol aaagagaaac actagccaaa cttggtgtaa gtgctgtacg tttcgttgag

CCAAATAATG CCATTACGGT TAATACACAA AACGAGTTTA CAACCAAACC

'01 ATCAAGTCAA GTGACAATTT CTGAAGGTAA GGCGÌGTTTC TCAAGTGGTA

ATGCCCACC ACTATGTACC AATGTTGCTG ACGATGCACA GCAG

### FIG.9A

1 ATGAACAAGA TATATCGTCT CAAATTCAGC AAACGCCTGA ATGCTTTGGT

TECTIGITETE GAATTEACAE GGGGTTGTGA CCATTECACA GAAAAAGGEA

101 GTGAAAAACC TGTTCGTACG AAAGTACGCC ACTTGGCGTT AAAGCCACTT

TCCGCTATAT TGCTATCTTT GGGCATGGCA TCCATTCCGC AATCTGTTTT

201 AGCGAGCGGT TTACAGGGAA TCAGCCTCGT ACACGGTACA GCAACCATGC

AAGTAGACGG CAATAAAACC ACTATCCGTA ATAGCGTCAA TGCTATCATC

301 AATTGGAAAC AATTTAACAT TGACCAAAAT GAAATGGTGC AGTTTTTACA

AGAAAGCAGC AACTCTGCCG TITITCAACCG TGTTACATCT GACCAAATCT &

401 CCCAATTAAA AGGGATTTTA GATTCTAACG GACAAGTCTT TTTAATCAAC

CCAAATGGTA TCACAATAGG TAAAGACGCA ATTAITTAACA CTAATGGCTT

501 TACTGCTTCT ACGCTAGACA TITICTAACGA AAACATCAAG GCGCGTAATT

TCACCCTTGA GCAAACCAAG GATAAAGCAC TCGCTGAAAT CGTGAATCAC

601 GGITTIAATTA CCGTTGGTAA AGACGGTAGC GTAAACCTTA TTGGTGGCAA

AGTGAAAAAC GAGGGCGTGA TTAGCGTAAA TGGCGGTAGT ATTTTCTTTAC

701 TTGCAGGGCA AAAAATCACC ATCAGCGATA TAATAAATCC AACCATCACT

TACAGCATTG CTGCACCTGA AAACGAAGCG ATCAATCTGG GCGATATTTT

## FIG.9B

801 TGCCAAAGGT GGTAACATTA ATGTCCGCGC TGCCACTATT CGCAATAAAG

GTAAACTTTC TGCCGACTCT GTAAGCAAAG ATAAAAGTGG TAACATTGTT

901 CICTCIGCCA AAGAAGGIGA AGCGGAAAIT GGCGGIGIAA ITICCGCICA

AAATCAGCAA GCCAAAGGTG GTAAGTTGAT GATTACAGGT GATAAAGTCA

1001 CATTAAAAAC AGGTGCAGTT ATCCACCTTT CAGGTAAAGA AGGGGGAGAG

ACTTATCTTG GCGGTGATGA GCGTGGCGAA GGTAAAAATG GTATTCAATT

TAAT GTATCAGGCA AAGAAAAAGG CGGGCGCCT ATTGTATGGG GCGATATTGC ATTAATTAAT S S 1101 AGCGAAGAAA ACCTCTTTAG AAAAAGGCTC GACAATTAAT GTATCAGGCA

1201 GGTAACATTA ATGCTCAAGG TAGCGATATT GCTAAAACTG GCGGCTTTGT

GGAAACATCA GGACATGACT TATCCATTGG TGATGATGTG ATTGTTGACG

1301 CTAAAGAGTG GTTATTAGAC CCAGATGATG TGTCCATTGA AACTCTTACA

TCTGGACGCA ATAATACCGG CGAAAACCAA GGATATACAA CAGGAGATGG

:401 gactaaagag tcacctaaag gtaatagtat ttctaaacct acattaacaa

ACTCAACTOT TGAGGAAATO CTAAGAAGAG GTTOTTATGT TAATATCACT

501 GCTAATAATA GAATTTATGT TAATAGCTCC ATCAACTTAT CTAATGGCAG

TITIPACACIT CACACTAAAC GAGATGGAGT TAAAATTAAC GGTGATATTA

### FIG.9C

1601 CCTCAAACGA AAATGGTAAT ITAACCAITA AAGCAGGCTC TIGGGTTGAT

GITCATAAAA ACATCACGCT TGGTACGGGT TITTTGAATA TIGTCGCTGG

1701 GGAITCIGTA GCTITTGAGA GAGAGGGCGA TAAAGCACGT AACGCAACAG

ATGCTCAAAT TACCGCACAA GGGACGATAA CCGTCAATAA ACATGATAAA

1801 CAATTTAGAT TCAATAATGT ATCTATTAAC GGGACGGGCA AGGGTTTAAA

GITTATIGCA AATCAAAATA ATTICACICA TAAATTIGAT GGCGAAATTA

1901 ACATATCTGG AATAGTAACA ATTAACCAAA CCACGAAAAA AGATGTTAAA

TACTGGAATG CATCAAAAGA CTCTTACTGG AATGTTTCTT CTCTTACTTT & 8

2001 GAATACGGTG CAAAAATTTA CCTTTATAAA ATTCGTTGAT AGCGGCTCAA

ATTCCCAAGA TTTGAGGTCA TCACGTAGAA GTTTTGCAGG CGTACATTTT

101 AACGGCATCG GAGGCAAAAC AAACTTCAAC ATCGGAGCTA ACGCAAAAGC

CTTATTTAAA TTAAAACCAA ACGCCGCTAC AGACCCAAAA AAAGAATTAC

201 CTATTACTTT TAACGCCAAC ATTACAGCTA CCGGTAACAG TGATAGCTCT

GIGATGITIG ACATACACGC CAATCITACC ICTAGAGCTG CCGGCATAAA

301 CATGGATTCA ATTAACATTA CCGGCGGCT TGACTTTTCC ATAACATCCC

ATAATCGCAA TAGTAATGCT TITGAAATCA AAAAAGACTT AACTATAAAT

## FIG.9D

2401 GCAACTGGCT CGAATTTTAG TCTTAAGCAA ACGAAAGATT CTTTTTATAA

TGAATACAGC AAACACGCCA TTAACTCAAG TCATAATCTA ACCATTCTTG

2501 GCGGCAATGT CACTCTAGGT GGGGAAAATT CAAGCAGTAG CATTACGGGC

AATATCAATA TCACCAATAA AGCAAATGTT ACATTACAAG CTGACACCAG

2601 CAACAGCAAC ACAGGCTTGA AGAAAAGAAC TCTAACTCTT GGCAATATAT

CIGITGAGGG GAATITIAAGC CTAACTGGTG CAAATGCAAA CATIGICGGC

2701 AATCTTTCTA TTGCAGAAGA TTCCACATTT AAAGGAGAAG CCAGTGACAA

CCTAAACATC ACCCCCACCT TTACCAACAA CGGTACCGCC AACATTAATA

2801 TAAAACAAGG AGTGGTAAAA CTCCAAGGCG ATATTATCAA TAAAGGTGGT

TTAAATATCA CTACTAACGC CTCAGGCACT CAAAAAACCA TTATTAACGG

54/82

1901 AAATATAACT AACGAAAAAG GCGACTTAAA CATCAAGAAT ATTAAAGCCG

ACCCCGAAAT CCAAATTGGC GGCAATATCT CACAAAAGA AGGCAATCTC

.001 ACAATTTCTT CTGATAAAGT AAATATTACC AATCAGATAA CAATCAAAAGC

AGGOGTTGAA GGGGGGCGTT CTGATTCAAG TGAGGCAGAA AATGCTAACC

101 TAACTATTCA AACCAAAGAG TTAAAAITTGG CAGGAGACCT AAATATTTCA

GGCTTTAATA AAGCAGAAAT TACAGCTAAA AATGGCAGTG ATTTAACTAT

201 TGGCAATGCT AGCGGTGGTA ATGCTGATGC TAAAAAAGTG ACTTTTGACA

AGGITAAAGA TICAAAAATC TCGACIGAGG GICACAAIGI AACACTAAAT

### FIG.9E

3301 AGCGAAGTGA AAACGTCTAA TOGTAGTAGC AATGCTGGTA ATGATAACAG

CACCGGITTA ACCATTTCCG CAAAAGATGT AACGGTAAAC AATAACGTTA

3401 CCTCCCACAA GACAATAAAT ATCTCTGCCG CAGCAGGAAA TGTAACAACC

AAAGAAGGCA CAACTATCAA TGCAACCACA GGCAGCGTGG AAGTAACTGC

CAGCAACAGA AAATCTTGTT ACCACAGAGA ATGCTGTCAT TAATGCAAGC 3501 TCAAAATGGT ACAATTAAAG GCAACATTAC CTCGCAAAAT GTAACAGTGA

3601 AGCGGCACAG TAAACATTAG TACAAAAQA GGGGATATTA AAGGTGGAAT

TGAATCAACT TCCGGTAATG TAAATATTAC AGCGAGCGGC AATACACTTA

55/82

3701 AGGTAAGTAA TATCACTGGT CAAGATGTAA CAGTAACAGC GGATGCAGGA

GCCTTGACAA CTACAGCAGG CTCAACCATT AGTGCGACAA CAGGCAATGC

801 AAATATTACA ACCAAAACAG GTGATATCAA CGGTAAAGTT GAATCCAGCT

CCGGCTCTGT AACACTTGTT GCAACTGGAG CAACTCTTGC TGTAGGTAAT

.901 ATTICAGGTA ACACTGTTAC TATTACTGCG GATAGCGGTA AATTAACCTC

CACAGTAGGT TCTACAATTA ATGGGACTAA TAGTGTAACC ACCTCAAGCC

001 AATCAGGCGA TATTGAAGGT ACAATTTCTG GTAATACAGT AAATGTTACA

GCAAGCACTG GTGATTTAAC TATTTGGAAAT AGTGCAAAAG TTGAAGCCAA

## FIG.9F

4101 AAATGGAGCT GCAACCTTAA CTGCTGAATC AGGCAAATTA ACCACCCAAA

CAGGCTCTAG CATTACCTCA AGCAATGGTC AGACAACTCT TACAGCCAAG

4201 GATAGCAGTA TCGCAGGAAA CATTAATGCT GCTAATGTGA CGTTAAATAC

CACAGGCACT TTAACTACTA CAGGGGATTC AAAGATTAAC GCAACCAGTG

1301 GTACCTTAAC AATCAATGCA AAAGATGCCA AATTAGATGG TGCTGCATCA

GGTGACCGCA CAGTAGTAAA TGCAACTAAC GCAAGTGGCT CTGGTAACGT

401 GACTGCGAAA ACCTCAAGCA GCGTGAATAT CACCGGGGAT TTAAACACAA

TAAATGGGTT AAATATCATT TCGGAAAATG GTAGAAACAC TGTGCCCTTA 98

501 AGAGGCAAGG AAATTGATGT GAAATATATC CAACCAGGTG TAGCAAGCGT

AGAAGAGGTA ATTGAAGCGA AACGCGTCCT TGAGAAGGTA AAACATTTAT

601 CTGATGAAGA AAGAGAAACA CTAGCCAAAC TTGGTGTAAG TGCTGTACGT

TICCTICAGC CAAATAATGC CATTACGGIT AATACACAAA ACGAGITTAC

101 AACCAAACCA TCAAGTCAAG TGACAATTTC TGAAGGTAAG GCGTGTTTCT

CAAGTOSTAA TGGCGCACGA GTATGTACCA ATGTTGCTGA CGATOGACAG

101 CAG

FIG.10A

COMPARISON OF DERIVED AMINO ACID SEQUENCE

50 KVRHLALKPL	KVRHLALKPL	KVRHLALKPL	KVRHLALKPL 5	100 TIRNSVNAII	TIRNSVNAII	TIRNSVNALI	TIRNSVNATT
EKGSEKPVRT	EKGSEKPVRT	EKGSEKPVRT	EKGSEKPVRT	ATMOVEGNKT	ATMOVECINKT	ATMOVIDGNKT	ATMOVECINKT
ELTRGCDHST	ELTRGCDHST	ELTRGCDHST	ELTRGCDHST	LQGMSVVHGT	LQGMSWHGT	LQGMSWHGT	LQGMSVVHGT
KRLNALVAVS	KRLNALVAVS	KRLNALVAVS	KRLNALVAVS	SIPQSVLASG	SIPOSVLASG	SIPQSVLASG	SIPQSVLASG
1 MNKIYRLKFS	MIKIYRLKFS	MIKIYRLKFS	MIKIYRLKFS	51 SAILLSLGMA	SAILLSLGMA	SAMLLSLGVT	SAMLLSLGVT
Hmv3com	Hrw4com	Hmv1com	Hrw2com	Hmw3com	· Hmw4com	Hmw2com	Hmw2com

### FIG. 10

				58	8/8	2				
	150 DSNSQVFLIN	DSNSQVFLIN	DSNSQVFLIN	DSNSQVFLIN		200	DKALAEIVNH	DKALAETVNH	DKALAEIVNH	DKALAEIVNH
	DQISQLKSIL	DŽISŽIKGIT	NQISQLKGIL	NQISQLKGIL			ARNFTI_EQTK	ARNFTLEQTK	ARNETTLEQTK	ARNFTLEQTK
	NSAVFNRVTS	NSAVFNRVTS	NSAVFNRVTS	NSAVFINRVTS		} { } }	TITOISNENIK	TLDISNENIK	TLDISNENIK	TLDISNENIK
	EMEQFIQESS	EMEQFIQESS	EMVQFTQENIN	EMVQFILQENIN			TTINTINZFILAS	IININSFTAS	IININSFTAS	IININSFTIĄS
101	NWKQFNIDQN	NWKQFINIDQN	NWKQFNIDQN	NWKQFNIDQN		151 PNGTTTCKDA		PNGITIGKDA	PNGITIGKDA	PNGITIGKDA
	Hrw3com	Hmv4com	Himlcom	Hmv2com		Hmw3com		Hmw4com	Hmw1com	Hmw2com

### FIG. 10(

	250 ISDIINPTIT	ISDIINPTIT	ISDIINPTIT	59/8 LILANIIQSI	300 VSKDKSGNIV
	ISLLAGOKIT	ISLLAGQKIT	ISLLAGOKIT	ISLLAGOKIT	300 RNKGKLSADS VSKDKSGNIV
	EGVISVNGGS	EGVISVNGGS	EGVISVNGGS	EGVISVNGGS	GNIINVRAATI
	GLITVGKDGS VNLIGGKVKN	VNLIGGKVKN	VNLIGGKVKN	GLITVGKDGS VNLIGGKVKN	INLGDIFAKG GNINVRAATI
201	GLITVGKDGS	GLITVGKDGS	GLITVGKDGS	GLITVGKDGS	251 YSIAAPENEA
	Hm3com	Hmv4com	Hmv1com	Hmw2com	Hm3com

# FIG. 10D.

VSKDKSGNTV	VSKDKCONITY	VSKDKSGNIV	
RNKGKLSADS	RNKGKI, SADS	RNKGKLSADS	
GNINVRAATI	GNINVRAATI	GNINVRAATI	
INLGDIFAKG	VNLGDIFAKG	VNLGDIFAKG	
YSIAAPENEA INLGDIFAKG GNINVRAATI RNKGKLSADS VSKDKSGNTV	YSIAAPENEA VNLGDIFAKG GNINVRAATI RNKGKI,SADG VSKDKSCNIV	YSIAAPENEA VNLGDIFAKG GNINVRAATI RNKGKLSADS VSKDKSGNIV	
Hmw4com	Hmw1com	Hmw2com	

IDLSGKEGGE 99/09 350 GGVISAQNQQ AKGGKLMITG DKVTLKTGAV IDLSGKEGGE IDLSGKEGGE IDLSGKEGGE DKVTLKTGAV DKVTLKTGAV DKVTLKTGAV GGVISAQNQQ AKGGKLMITG GGVISAQNQQ AKGGKLMITG GGVISAQNQQ AKGGKLMITG LSAKEGEAEI LSAKEGEAEI LSAKEGEAEI LSAKEGEAEI 301 Hmw3com Hmw4com Hmw1com Hmw2com

400 IVWGDIALID IVWGDIALID IVWGDIALID IVWGDIALID TYLGGDERGE GKNGIQLAKK TTLEKGSTIN VSGKEKGGRA TTLEKGSTIN VSGKEKGGRA TTLEKGSTIN VSGKEKGGRA GKNGIQLAKK TTLEKGSTIN VSGKEKGGRA GKNGIQLAKK GKNGIQLAKK TYLGGDERGE TYLGGDERGE TYLGGDERGE 351 Hmw3com Hmw4com Hmw2com Hmw1com

# FIG. 10E.

401

450 SGHYLSIDDN AIVKTKEWLL DPENVTIEAP DPDDVSIETL DPDNVTINAE DPDDVTIEAE VIVDAKEWLL AIVKTKEWLL AIVDAKEWLL SGHDLSIGDD SGHYLSIESN SGHDLFIKDN IAKTGGFVET IAKTGGFVET IAKTGGFVET IAKTGGFVET GNINAQGK.D GNINAQGS.D GNINAQGSGD GNINAQGSGD Hmw3com Hmw4com Hmw1com Hmw2com

451

61/82 500 ILRRGSYVNI SASRVELGAD RNSHSAEVIK VTLKKNNTSL TTLTNTTISN LLKSAHVVNI ILKKGTFVNI TTLTNTTISN YLKNAWTMNI PTLTNSTLEQ TTLTNTTLES SDPKKNSELK ESPKGNSISK STPKRNKE.K QGYTTGDGTK DEYTGSGNSA DEFPTGTGEA TSGRNNTGEN TAGRSNTSED DPLRNNTGIN Hmw3com Hmw4com Hmw1com Hmw2com

501

550 .E...GGNLT NE...NGNLT GVEINNDITT GDDTRGANLT · · · SKGGNLT SISIERGSHL ILHSEGQGGQ GVQIDKDITS GVKINGDITS ILHSKGQRGG GVQIDGDIT. TLHTK...RD TLWSEGRSGG SINLSNGS.L SINL. SNGSL SINGSNGSHL TARRKLTVNS TANNRIYVNS TANQRIYVNS TASRKLTVNS Hmw3com Hmw4com Hmw1com Hmw2com

· · LTHNLSGT

# FIG. 10F

551

009 ..NNLTITAQ AFEREGDKAR NATDAQITAQ DANNLTITAQ KNITLD.QGF LNITA.AS.V AFEGGNNKAR IYSGGWVDVH KNITLGS.GF LNITTKEGDI AFEDKSGR.. AFEKGSNQV. LNIVAGDS.V KNISLGAQGN INITAKQD.I KNITLGT.GF IKAGSWVDVH  ${ t IYSGGWVDVH}$ IYSGGWVDVH Hmw3com Hmw4com Hmw1com Hmw2com

62/82 650 GNISNKFDGT .NFTHKFDGE YAITNKFEGT GTITSG.NSN GFRFNNVSLN SLGGKLSFTD SREDRGRRTK GTITVNKDDK QFRFNNVSIN GTGKGLKFIA NQN..... KRTN....K SVNN GFRFNNVSLN GTGSGLQFTT GTGKGLNIIS DFRANNVSLN GTIT. SGNQK GTVTITGEGK 601 Hmw3com Hmw4com Hmw1com Hmw2com

700 LNISGTVDIS MKAPKVSWFY RD.KGRTYWN VTTLNVTSGS KFNLSIDSTG KFTF.IKFVD EFNLTIDSRG VSSLTLNTVQ LTSLNVSESG QTTKKDVKYW NA.SKDSYWN MVLPKNESGY DKFKGRTYWN INISGIVTIN LNISGKVNIS 651 Hmw3com Hmw1com Hmw4com

FIG. 10G.

QTTRKNTSYW QTSHD.SHWN VSALNLETGA NFTF.IKYIS INISGNITIN Hmw2com

750 NFSIKASIMP LFKLKPNAAT NFDIKAPIGI SFNLKEGAKV NFKLKPNENM TFNIAQGSTA NFNIGANAKA TFNVERNARV ITFN....KA V..N.;.GNM VHFNGIGGKT ISFN...KDT RSSAGVNFNG IRNA..ELNG LRSSRRSFAG · · · · PYNLNG SNSKGLTTQY SGSNS...QD SGSTG...PS SDSAGTLTQ. 701 Hmw3com Hmw4com Hmw1com Hmw2com

63/82 0 0 8 GGSVNFKLN ASSSNIQTPG VIIKSQNFNV AGINMDSINI GGSVDFTLL ASSSNVQTPG VVINSKYFNV AELKMSEINI SDSSVMFDIH A...NLTSRA .GGSVFFDIY ANHS...GRG FNANITATGN FNEDISVSG. FLANITATG. FNGNISVSG. DPKKELPIT. FKSNANYAL. NKYSSLNYAS NTSKPLPI.R 751 Hmw3com Hmw4com Hmw1com Hmw2com

850 ENDLNLNATG GNITIRQVEG T.. DSRVNKG SFYNEYSKHA SNFSLKQTKD KKDLTINATG EGSTETAFSI HNRNSNAFEI SGGSTLNLKA TGGLDFSITS 801 Hmw3com Hmw4com

TRDTLNITGN

ISESATFKGK

RDRVIKLGSL LVNGSLSLTG ENADIKGNLT

Hmw2com

FIG. 10H

T..DGMIGKG DFYDGYARNA EKDLTLNATG GNITLLQVEG SNFSLRQTKD NKDLTINATN SGSTKTGFSI HVRGDDAFKI SNGANFTLNS STGSSLRFKT Hmw1com Hmw2com

900 851

ADTSNSNTGL VAAKKNITFK GGNITFGSQK ATTEIKGNVT INKNTNATLR GANFAEN... GSDFDNHQ. ITNKANVTLQ INNNANVTLI SSSSITGNIN AVTEIEGNVT GGNITFGSRK GGNVTLGGEN INSSHNLTIL IVAKKNITFE Hmw3com Hmw4com Hmw1com

64/82 ANNAPNQQNI IEKAANVTLE INSTYNISIL GGNVTLGGQN SSSSITGNIT Hmw2com

950 901

INNGNLTTAG SIINIAGNLT VSKGANLQAI TNYTFNVAGS ASDNLNITGT TNFTFNVGGL IAEDSTFKGE VESNANFKAI INSGNLTAGG NIVNIAGNLT SVEGNLSLTG ANANIVGNLS KSPLNIAGNV KKRTLTLGNI KPLTIKKDVI Hmw3com Hmw1com Hmw4com

1000 951

FIG. 10I

TTNSDTTYRT IIKGNISNKS IINGNITNEK IISGNITNKN IIGGDIINNK TTHAKRNQRS TTNASGTQKT TTNSSSTYRT DINNKGGLNI DINNISSINI DIDNSKNLSI ITQGVVKLG. NVTNDGDLNI IARGGAKFK. IAKGGARFK. IKQGVVKLQG FDNNGASNIS FDNKGNSNIS FINNGTAEIN FTNNGTANIN Hmw3com Hmw4com Hmw1com Hmw2com

1050 1001

TIKAGVDGEN 8/59 SDKVNITNQI TIKAGVEGGR TIKAGVEGGR SDKINITKQI TIKKGIDGED SDKVNITNQI SDKINITKQI SQKEGNLTIS SQKEGNLTIS SQKEGNLTIS SQKEGNLTIS DAEIQIGGNI DTEMQIGGDI DAEIQIGGNI GSLNITDSNN DAEIQIGGNI GDLNIIDKKS GDLNIKNIKA GDLNITNEGS Hmw3com Hmw4com Hmw1com Hmw2com

1100 1051

DLTIGNASGG DLTIGNASGG DLTIGNTNSA DLTIGNSNDG SDSSEAENAN LTIQTKELKL AGDLNISGFN KAEITAKNGS TEDLSISGFN KAEITAKDGR KAEITAKNGS KAEITAKDGS TQDLNISGFN AGDLNISGFN LTIQTKELKL SDSDATNNAN LTIKTKELKL LTIKTKELKL SDSSEAENAN SSSDATSNAN Hmw1com Hmw3com Hmw4com Hmw2com

NGKASIT.

# FIG. 10J

1101

1150 SNAGNDNSTG SNAGNDNSTG TEDSSDNNAG RESNSDNDTG EVKT..SNGS EVKT..SNGS KVETSGSNNN KVKTSSSNGG TDGHNVTLNS ADGHNVTLNS TDGHNVTLNS ADGHKVTLHS FDKVKDSKIS FNQVKDSKIS FNNVKDSKIS FDKVKDSKIS N..ADAKKVT N..ADAKKVT D.GTNAKKVT NSGAEAKKVT Hmw3com Hmw4com Hmw1com Hmw2com

1200 1151

66/82 TGSVEVTAQN TKEGTTINAT TGSVEVTAQN TGNVEIT.. TKEGTTINAT TKTGTTINAT TTAGSTINAT LTISAKDVTV NNNVTSHKTI NISAAAGNVT SISATSGEIT NITA.SEKVT NISAAAGNVT NNNITSHKAV LTISAKDVTV NNNVTSHKTI LTITAKNVEV NKDVTSLKTV LTIDAKNVTV Hmw3com Hmw4com Hmw1com Hmw2com

1250 1201

TGDIKGGIES TGDIKGGIES TGDIKGGIES TSGTVNISTK TSGTVNISTK ......AQ GTIKGNITSQ NVTVTATENL VTTENAVINA GTIKGNITSQ NVTVTATENL VTTENAVINA Hmw3com Hmw4com Hmw1com

	_					67/	32						
	. 1300	ISATTGNANI	ISATTGNANI	IKG.TESVTT		1350	ADSGKLTSTV	ADSGKLTSTV	ATESLTTQSN	ATVDLTTKSG	1400	NSAKVEAKNG	NSAKVEAKNG
TK		GQDVTVTADA GALTTTAGST	GALTTTAGST	GALTTLAGST			NISGNTVTIT	NISGNTVTIT	TISGGTVEVK	TISGNTVSVS		TASTGDLTIG	TASTGDLTIG
		GQDVTVTADA	GQDVTVTADA	GNTVTVTANS			VATGATLAVG	VATGATLAVG		· · · · · · · · · · · · · · · · · · ·		GTISGNTVNV	GTISGNTVNV
: : : : : :		GNTLKVSNIT	GNTLKVSNIT	EGALAVSNIS			VESSSGSVTL	VESSSGSVTL			-	TTSSQSGDIE	TTSSQSGDIE
	T C フ T	TSGNVNİTAS	TSGNVNITAS	SSGSVTLTAT	•	1301	TTKTGDINGK	TTKTGDINGK	SSQSGDIG	GDIS	1351	GSTINGTNSV	GSTINGTNSV
FIG.10K. Hmw2com		Hmw3com	Hmw4com	Hmw1com	Hmw2com		Hmw3com	Hmw4com	Hmw1com	Hmw2com		Hmw3com	Hmw4com

GTISGNTVNV TANAGDLTVG NGAEINATEG NGAEINATEG TANAGDLTVG GTISGNTVNV SKIKATTGEA NVTSATGTIG SKIEAKSGEA NVTSATGTIG Hmw1com Hmw2com

1401

68/82 1450 SSNGQTTLTA KDSSIAGNIN AANVTLNTTG AANVTLNTTG AANVTLNTTG STKGQVDLLA QNSSIAGNIN AANVTLNTTG KDSSIAGNIN QDSSVAGSIN SSNGQTTLTA SAKGQVNLSA AATLTAESGK LTTQTGSSIT LTTEASSHIT AATLTAESGK LTTQTGSSIT LTTEAGSSIT AATLTTSSGK AATLTATGNT Hmw3com Hmw4com Hmw1com Hmw2com

1500

TLTTTGDSKI NATSGTLTIN AKDAKLDGAA SGDRTVVNAT NASGSGNVTA NASGSGNVTA NANGSGSVIA SGDSTEVNAV NASGSGSVTA LGNHTVVNAT SGDRTVVNAT AKDAKLDGAA AKDAELNGAA AKDAKLNGDA NATSGTLTIN KATSGTLTIN NATSGTLTIN TLTTVAGSDI TLTTTGDSKI TLTTVKGSNI Hmw3com Hmw4com Hmw1com Hmw2com

1451

1550

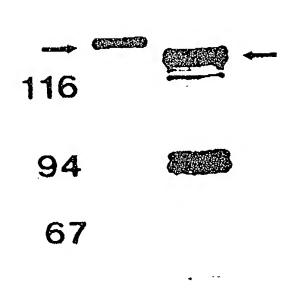
	_
	Σ
	_
(	$\supset$
	⊣
•	
	•
_	_
•	り
_	_
H	
r.,	_
ſΞ	4

, ,	LQPGVASVEE	DLITINGINI ISENGRNTVR LRGKEIDVKY IQPGVASVEE		LVPGIASVDE	QPGVASVEE	
DLNTINGLNI ISENGRNTVR 1.PCVETDXXXX +0	TYOUTHOURE	LRGKEIDVKY .	DLITINGLNI ISKNGINTVI, I,KCVKIDVWV 1000		L ISKDGRNTVR LRGKEIEVKY IQPGVASVEE	
ISENGRNTVR	T. ERIKA OIKA O T	LOENGKIN'I'VR	ISKNGINTVI,	i c	ISKUGKNTVR	
DLNTINGLNI	DENTINGIAL		DLITINGLNI	DI,NTVMCT NI		
KTSSSVNITG	KTSSSVNITG		PI.TNIARCE T	ATSSSVNITG		
Hmw3com	Hmw4com	Hmw1 com	)	Hmw2com		

1600	VNTQNEFTTK	K VKNIGAGAGA SAV KFVEPNNAIT VNTQNEFTTK 9	VDTQNEFATR 8	VNTQNEFTTR
VKDLSDEERE TLAKLGVSAV PRVERNARATER	TI OEFINALI	RFVEPNNAIT	TITEL ALAKLGVSAV RFIEPNNTIT VDTQNEFATR	VKDLSDEERE TLAKLGVSAV RFVEPNNTIT VNTQNEFTTR
TLAKLGUSAV	TI,AKI,GWAW		ALAKLGVSAV	TLAKLGVSAV
VKDLSDEERE	VKDLSDEERE	UKDI SDEBE	באקקייטטטייטיי	VKDLSDEERE
VIEAKRVLEK	VIEAKRVLEK	VIEAKRILEK	7. THAND 11.	Y TONKY LEK
Hmw3com	Hmw4com	Hmw1com	Hmw2com	

(	1632 3 QQ		א א ע	ŎŎ S		X		AQ 5	
	1632 PSSQVTISEG KACFSSGNGA RVCTNVADDG OO			PSSQVTISEG KACFSSGNGA RVCTNVADDG QQ		FLSKIVISEG RACFSNSDGA TVCVNIADNG R.		RVCTNVADDO	
		KACFSSGNGA		NACFSSGNGA		KACFSNSDGA		KACFSSGNGA	
TOOT		FSSCV111SEG	DSSOMME DESCRIPTIONS OF THE PROPERTY OF THE PR	クロストトンド・トンドゥ	DT. ATTT GD. TG	THOILY LOEG	D00011110000	FSSKVIISEG KACFSSGNGA RVCTNVADDG OF	
-0	Hmw3com		Hmw4com		Hmw1com		Hmw2com		

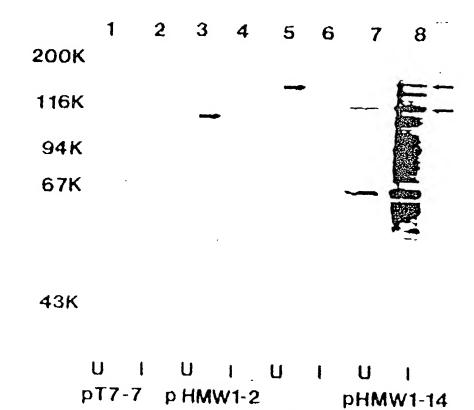
kDa 200



#### 43 HMW1

FIG. 11 HMW 2

WESTERN IMMUNOBLOT ASSAY OF PHAGE LYSATES CONTAINING EITHER THE HMW1 OR HMW2 RECOMBINANT PROTEINS. LYSATES WERE PROBED WITH AN *E. COLI-* ABSORBED ADULT SERUM SAMPLE WITH HIGH-TITER ANTIBODY AGAINST HIGH-MOLECULAR-WEIGHT PROTEINS. THE ARROWS INDICATE THE MAJOR IMMUNOREACTIVE PROTEIN BANDS OF 125 AND 120 kDa IN THE HMW1 AND HMW2 LYSATES, RESPECTIVELY.



#### FIG. 12

WESTERN IMMUNOBLOT ASSAY OF CELL SONICATES PREPARED FROM E. COLI TRANSFORMED WITH PLASMID pT7-7 (LANES 1 AND 2) pHMW1-2 (LANES 3 AND 4), pHMW1-4 (LANES 5 AND 6), OR pHMW1-14 (LANES 7 AND 8). THE SONICATES WERE PROBED WITH AN E. COLI-ABSORBED ADULT SERUM SAMPLE WITH HIGH -TITER ANTIBODY AGAINST HIGH - MOLECULAR -WEIGHT PROTEINS. LANES LABELED U AND I REPRESENT SONICATES PREPARED BEFORE AND AFTER INDUCTION OF THE GROWING SAMPLES WITH IPTG, RESPECTIVELY. THE ARROWS INDICATE PROTEIN BANDS OF INTEREST AS DESCRIBED IN THE TEXT.

**PHMW1-4** 

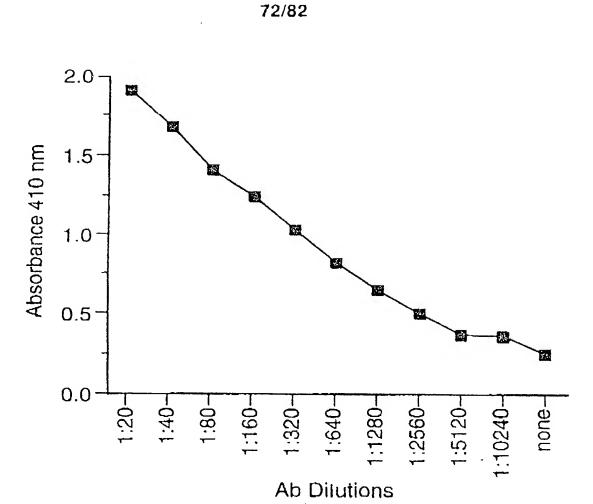


FIG. 13
ELISA WITH rHMW1 ANTISERUM ASSAYED AGAINST PURIFIED FILAMENTOUS HEMAGLUTININ OF B. PERTUSSIS. Ab, ANTIBODY.

200K

116K

94K

67K

43K

5 7 12 14 15 16 17 18

#### FIG. 14

WESTERN IMMUNOBLOT ASSAY OF CELL SONICATES FROM A PANEL OF EPIDEMIOLOGICALLY UNRELATED NONTYPEABLE H. INFLUENZAE STRAINS. THE SONICATES WERE PROBED WITH RABBIT ANTISERUM PREPARED AGAINST HMW1-4 RECOMBINANT PROTEIN. THE STRAIN DESIGNATIONS ARE INDICATED BY THE NUMBERS BELOW EACH LANE.

200K

116K

94K

67K

43K

5 7 12 14 15 16 17 18

#### FIG. 15

WESTERN IMMUNOBLOT ASSAY OF CELL SONICATES FROM A PANEL OF EPIDEMIOLOGICALLY UNRELATED NONTYPEABLE H. INFLUENZAE STRAINS.THE SONICATES WERE PROBED WITH MONOCLONAL ANTIBODY X3C, A MURINE IgG ANTIBODY WHICH RECOGNIZES THE FILAMENTOUS HEMAGGLUTININ OF B. PERTUSSIS (13).THE STRAIN DESIGNATIONS ARE INDICATED BY THE NUMBERS BELOW EACH LANE.

1 2 3 4

kDa

200

116

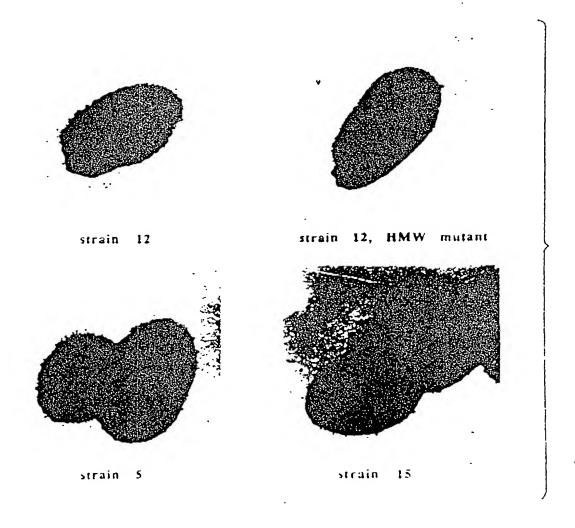
94

67

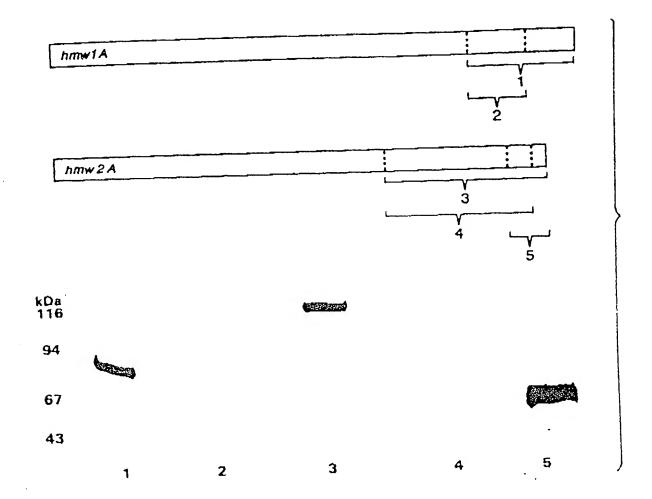
43

#### FIG. 16

IMMUNOBLOT ASSAY OF CELL SONICATES OF NONTYPABLE H. INFLUENZAE STRAIN 12 DERIVATIVES. THE SONICATES WERE PROBED WITH RABBIT ANTISERUM PREPARED AGAINST HMW-1 RECOMBINANT PROTEIN. LANES: 1, WILD-TYPE STRAIN; 2, HMW-2 MUTANT; 3, HMW-1 MUTANT; 4, HMW-1 / HMW-2 DOUBLE MUTANT.

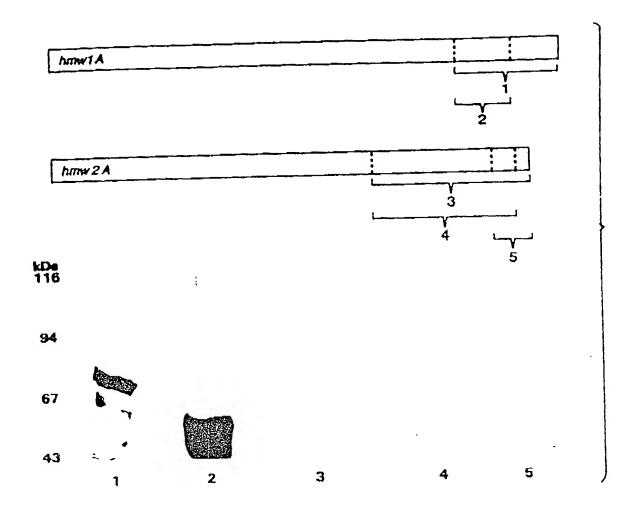


IMMUNOELECTRON MICROSCOPY WITH Mab AD6 FIG.20



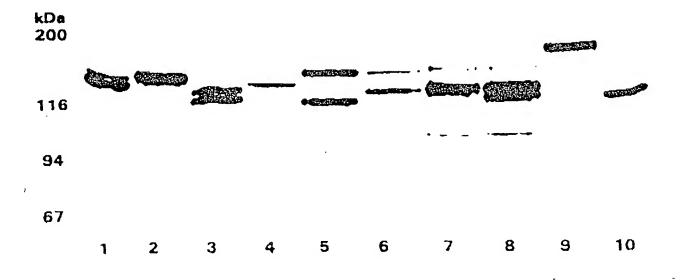
WESTERN IMMUNOBLOT ASSAY WITH Mab AD6 AND HMW1A OR HMW2A RECOMBINANT PROTEINS

FIG.21



WESTERN IMMUNOBLOT ASSAY WITH Mab 10C5 AND HMW1A OR HMW2A RECOMBINANT PROTEINS

**FIG.22** 



WESTERN IMMUNOBLOT ASSAY WITH Mab AD6 AND TEN UNRELATED NONTYPABLE HAEMOPHILUS INFLUENZAE

**FIG.23**